



FILOZOFICKÁ FAKULTA
Univerzita Karlova

Ústav anglického jazyka a didaktiky

Diplomová práce

Bc. Eliška Kubánková

Diskurzní konektory v textech nerodilých mluvčích angličtiny

Discourse connectives in texts written by non-native speakers of English

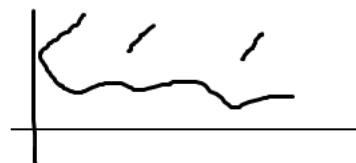
Praha 2018

Vedoucí práce: PhDr. Magdaléna Rysová, Ph.D.

Na tomto místě bych ráda poděkovala své vedoucí práce, PhDr. Magdaléně Rysové, Ph.D., za odborné vedení, vstřícnost a za poskytnutí cenných konzultací a připomínek k tématu. Poděkování patří též mé rodině, která mi byla od začátku studia velkou oporou.

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 30. dubna 2018

A handwritten signature in black ink, consisting of a stylized 'K' followed by a series of connected loops and a final horizontal stroke.

Bc. Eliška Kubánková

Souhlasím se zapůjčením diplomové práce ke studijním účelům. I have no objections to the MA thesis being borrowed and used for study purposes.

Abstrakt

Tématem práce je analýza diskurzních konektorů jako prostředků textové koheze ve formálních textech psaných nerodilými mluvčími angličtiny. Materiálem pro výzkum byl korpus 161 formálních dopisů psaných studenty, jejichž jazyková úroveň odpovídá úrovni B2 ve Společném evropském referenčním rámci pro jazyky. Hlavním cílem práce bylo zjistit, jaké typy konektorů se v textech nerodilých mluvčích vyskytují a s jakou frekvencí, a tím specifikovat úroveň B2 v oblasti psaného projevu. Zkoumanými jevy byly kromě frekvence druhy konektorů (jednoslovné/víceslovné, primární/sekundární), jejich morfologické a syntaktické kategorie, rozsah (srov. mezivětné či vnitrovětné konektory), pozice ve větě; dále typy diskurzních argumentů a sémantické typy vztahů, které konektory vyjadřovaly. Výsledky práce ukázaly, že konektory byly nejčastěji jednoslovné primární výrazy realizované spojkami a příslovečnými určeními. Nejčastěji používaným konektorem byla spojka *and*, druhým nejčastějším spojka *because*. Častý výskyt měla i adverbia *so* a *also*. Konektory nejběžněji vyjadřovaly vztahy konjunkce a příčiny, což mohlo být ovlivněno typem zkoumaných textů. V textech se objevily i některé sekundární konektory, jako například varianty vět *The reason is* a *That is why*. Konektory se častěji objevovaly na začátku argumentu a vyjadřovaly vnitrovětný vztah. Výzkum prokázal celkové nadužívání konektorů nerodilými mluvčími. Ve výsledcích se též projevil časté chyby v používání konektorů z hlediska jejich významu, pozice a stylistiky.

Klíčová slova: text, diskurz, diskurzní konektory, diskurzní vztahy, koheze, angličtina jako cizí jazyk, formální angličtina, nerodilí mluvčí

Abstract

The thesis discusses the topic of discourse connectives as means of textual cohesion in formal texts written by non-native speakers of English. The material used for the analysis is a corpus of 161 application emails written by students whose language proficiency corresponds to the level B2 in the Common European Framework of Reference. The aim of the thesis was to determine the frequency and kinds of connectives used by B2 speakers, and to help specify the level B2 in terms of writing. The analyzed aspects were frequency, types of connectives (one-word/multi-word, primary/secondary), their morphological and syntactic classes, scope (inter-/intra-sentential), position in the sentence; further the types of discourse arguments and semantic types of relations expressed. It was found that most connectives were one-word, primary expressions realized by conjunctions and adverbials. The most frequently occurring connective was *and*, followed by *because* and the adverbs *so* and *also*. The most frequent semantic type of relation expressed was expansion, followed by contingency (cause), which can be attributed to the text type. Some secondary structures were also found, including different modifications of the clauses *The reason is* and *That is why*. The preferred position of connectives was initial, and their scope was mostly intra-sentential. The thesis revealed a general overuse of connectives by B2 speakers. The results also indicated errors in the use of connectives in terms of their meaning, position and stylistic use.

Key words: text, discourse, discourse connectives, discourse relations, cohesion, English as a foreign language, formal English, non-native speakers

Table of Contents

1	INTRODUCTION	11
2	THEORETICAL BACKGROUND.....	14
2.1	DISCOURSE.....	14
2.1.1	Approaches to Discourse, Discourse Analysis	14
2.1.1.1	The Formalist Approach to Discourse	14
2.1.1.2	The Functionalist Approach to Discourse	15
2.1.2	Text and Texture.....	15
2.1.3	Cohesion and Coherence	16
2.1.3.1	Cohesion in Halliday and Hasan (1976)	18
2.1.4	Lexical and Grammatical Cohesion.....	18
2.1.4.1	Lexical Cohesion: Reiteration, Collocation, General Nouns.....	19
2.1.4.2	Grammatical Cohesion: Substitution, Ellipsis, Reference.....	19
2.1.4.3	Grammatical Cohesion: Conjunction.....	20
2.1.4.4	Cohesion and Coherence in Evaluating Writing.....	21
2.1.5	Defining the Basic Unit of Discourse	21
2.1.5.1	The Penn Discourse Treebank and Discourse Arguments.....	22
2.1.5.2	Syntactic Features of Discourse Arguments	23
2.2	DISCOURSE RELATIONS.....	23
2.2.1	Explicit and Implicit Relations	24
2.2.2	External and Internal Relations.....	25
2.2.3	Semantic and Pragmatic Types of Relations	25
2.3	DISCOURSE CONNECTIVES	28
2.3.1	Terminology.....	28
2.3.2	Approaches to Discourse Connectives	29
2.3.2.1	Halliday and Hasan's (1976) Approach to Discourse Connectives...	29
2.3.2.2	Fraser's (1990, 1998) Approach to Discourse Connectives	30

2.3.2.3	Schiffrin's (1987) Approach to Discourse Connectives	30
2.3.3	Aspects of Definition	31
2.3.3.1	Word Class.....	31
2.3.3.2	Pragmatic Aspect	31
2.3.3.3	Other Aspects.....	32
2.3.4	Coherence and Relevance Theory	33
2.3.4.1	Coherence-based Approach	34
2.3.4.2	Relevance-based Approach.....	34
2.3.4.3	Connectives in Coherence and Relevance Theory	35
2.3.5	Types of Discourse Connectives.....	35
2.3.5.1	Primary and Secondary Connectives	36
2.3.5.2	Modified, Parallel, and Conjoined Connectives	36
2.3.5.3	Contrastive, Collateral and Inferential Connectives	37
2.4	ENGLISH AS A FOREIGN LANGUAGE.....	38
2.4.1	Studies on Discourse Connectives in Texts of Non-Native Speakers	38
2.4.2	The Common European Framework of Reference (CEFR)	41
2.4.3	B2 in the CEFR.....	42
2.4.3.1	Written Production in B2.....	43
2.4.3.2	Discourse Competence in B2.....	44
3	MATERIAL AND METHOD	46
3.1	Subjects.....	46
3.2	Material.....	47
3.3	Data Processing	48
3.4	Hypothesis	50
4	ANALYSIS.....	51
4.1	General Findings.....	51
4.2	Morphological and Syntactic Classes of Connectives.....	53
4.2.1	Conjunctions	54

4.2.1.1	Cohesive vs. Structural “And”	56
4.2.1.2	Subordinating “Because”	59
4.2.1.3	Correlative Pairs	60
4.2.1.4	Other Multiple Conjunctions	61
4.2.2	Adverbs - Conjuncts	63
4.2.3	Prepositions and Prepositional Phrases	65
4.2.4	Clauses	66
4.3	One-Word/Multi-Word, Primary/Secondary Connectives	68
4.4	Position in Sentence.....	71
4.4.1	Overview.....	71
4.4.2	Position and Scope (Intra-/Inter-Sentential)	74
4.4.3	Final Connectives	77
4.5	Discourse Arguments	77
4.5.1	Verbal Arguments.....	78
4.5.2	Nominal Arguments.....	79
4.6	Semantic and Pragmatic Types of Relations	80
4.6.1	Overview.....	80
4.6.2	Semantics of “And”	84
5	CONCLUSION.....	86
	References	92
	Résumé	97
	Appendix	100

List of Abbreviations

Arg	Argument
Arg1	Argument 1 in a Discourse Relation
Arg2	Argument 2 in a Discourse Relation
B2	Designation for the “Vantage” level in the CEFR
CEFR	Common European Framework of Reference
EFL	English as a Foreign Language
ESL	English as a Second Language
PDTB	Penn Discourse Treebank
SFL	Systemic Functional Linguistics
VŠE	Vysoká škola ekonomická (University of Economics in Prague)

List of Tables and Figures

Tables

Table 1: Basic Units of Discourse According to Johnstone (2002)	21
Table 2: Hierarchy of Abstract Objects According to Asher (1993: 57)	22
Table 3: Table of Semantic Relations According to Halliday and Hasan (1976: 242)	26
Table 4: Table of Reference Levels in the CEFR (2001)	42
Table 5: B2 Written Production in the CEFR (2001)	43
Table 6: Descriptive Scale of Cohesion and Coherence in the CEFR (2001)	44
Table 7: B2 Assignment – Erasmus Application Letter	47
Table 8: Example of B2 Text from the Student Corpus (Text No. 13)	49
Table 9: Example of Connective Annotation in Microsoft Excel (Text No. 13)	49
Table 10: Average Ratios of Connectives per Sentence and Sentences per Connective	52
Table 11: Initial Use of "Because " in B2 Texts	59
Table 12: Examples of Multi-Word Connectives Used by B2 Speakers	61
Table 13: Superfluous Use of Connectives	62
Table 14: Types of Conjuncts in the Texts by B2 Speakers	64
Table 15: List of Secondary Connectives Realized by Clauses	67
Table 16: List of Redundant Connective Expressions	70
Table 17: Position of Connectives in Sentence: Initial, Post-Initial, Medial, Final	72
Table 18: Position of the Connective "also" in the Texts by B2 Speakers	73
Table 19: Ratio of Inter- vs Intra-Sentential Connectives	74
Table 20: Initial Intra-Sentential Use of "If"	75
Table 21: "Since" used as Preposition	80
Table 22: Semantic Types of Relations: Expansion, Contingency, Comparison, Temporal ...	80
Table 23: Types of Semantic Relations and Corresponding Connectives in the B2 Texts	82
Table 24: Overview of Most Frequent Connectives and Semantic Relations They Express ...	83
Table 25: Subtypes of Expansion (Conjunction, List) as Defined by the PDTB (2007)	84
Table 26: Semantic Relations Expressed by "and" in the B2 Texts	84

Figures

Figure 1: Semantic Relations According to the PDTB (Prasad et al., 2007)	27
Figure 3: Overall Frequencies of Connectives per Text	51
Figure 4: Number of Texts with Different Connective Frequencies.....	53
Figure 5: Morphological and Syntactic Classes of Connectives.....	54
Figure 6: Ratio of Coordinating and Subordinating Conjunctions	54
Figure 7: Coordinating Conjunctions in B2 Texts	55
Figure 8: Subordinating Conjunctions in B2 Texts.....	55
Figure 9: Structural Use of "and" in B2 Texts	56
Figure 10: Example of B2 Text (No. 87) with Cohesive Use of "and"	57
Figure 11: List of Correlative Pairs (Parallel Connectives)	60
Figure 12: Ratio of One-Word and Multi-Word Connectives	69
Figure 13: Frequency of Primary and Secondary, One-word or Multi-Word Connectives	69
Figure 14: Position of Connectives in Sentence: Initial, Post-Initial, Medial, Final.....	72
Figure 15: Ratio of Semantic Types of Relations	81

1 INTRODUCTION

The topic of this thesis are discourse connectives in formal letters written by non-native speakers of English. Discourse connectives are linking units that signal semantic relations in text, such as conjunctions (*and, but, or*), adverbs (*moreover, however, therefore*), or some multi-word units (*under these conditions, to sum up, the reason is*). As such, discourse connectives play a fundamental role in the structuring of text and its coherence. The study aims to analyze the frequency and use of discourse connectives in texts written by speakers whose language proficiency corresponds to the level B2 in to the Common European Framework of Reference (CEFR), thus to contribute to the specification of the reference level B2 in terms of written production.

During the past ten years, the study of discourse connectives has become a key topic in linguistics. The term has different meanings for different groups of researchers, and the phenomenon can be found under a variety of labels, including *discourse markers, discourse particles, pragmatic markers, discourse operators, connectors, pragmatic operators, pragmatic connectives, phatic connectives, semantic conjuncts, conjunctive elements, linking adverbials, conjunctive adverbials*, and others. The fact that there is no clear-cut term to label the phenomenon points at the underlying complexity of the topic. Most researchers agree that discourse connectives are elements that relate discourse segments. However, there is no agreement on how exactly discourse connectives should be defined, and what specific functions they have in text. Some of the most extensive studies on discourse and discourse connectives in English include works by Halliday and Hasan (1976), Schiffrin (1987, 1994 2001), Fraser (1988, 1990, 1996, 1999, 2009) and Blakemore (1987, 2002). Although authors use different terms for the concept, their studies all aim to explain the meaning of discourse connectives, the functions they manifest, and the ways they pattern in text.

To acquaint the reader with the subject matter, we will first provide an overview of the main approaches to discourse, discourse analysis and discourse connectives. All key terms within each topic will be explained and defined. We will first present two fundamental approaches to discourse – the functionalist and formalist approach. Further, we will explain the concepts of cohesion and coherence and their relation to the studied phenomenon. This part will be largely based on the work *Cohesion in English* by Halliday and Hasan (1976), which provides a comprehensive framework for the analysis of grammatical and lexical cohesive devices in text. To understand what kinds of segments connectives join, we will introduce approaches to the basic unit of discourse. Our definition will be based on the Penn Discourse

Treebank (PDTB, Prasad et al., 2007), a large corpus that is annotated with information related to discourse structure and semantics and focused on encoding discourse relations. The PDTB (Prasad et al., 2007) will also be the main source for the classification of discourse relations, including their semantic and pragmatic types. A significant portion of the theoretical part will be dedicated to the definition of discourse connectives. We will look at some of the main approaches to connectives, including the definitions within coherence and relevance theory, two basic theories that take a slightly different perspective on discourse connectives and their role in text. Then, we will explain some of the ways discourse connectives are classified.

Discourse connectives have become an increasingly popular topic among linguists, with many articles appearing yearly. Discourse connectives have been analyzed in a number of languages, including Arabic (Al-Saif and Markert, 2010), Chinese (Xue, 2005), Turkish (Zeyrek and Weber, 2008), or Czech (Rysová, 2015; Zikánová et al., 2015). We will present some recent studies on the use of discourse connectives in texts by non-native speakers of English (Aysu, 2017; Chen, 2006; Granger and Tyson, 1996; Lahuerta Martinez, 2004; Shaw, 2009; Tomešová, 2017), which will be later compared with the results of our own study.

Since this study will concern the language of non-native speakers, specifically of level B2 speakers as defined by the CEFR, we will introduce the CEFR and its reference levels, focusing on the level B2 and the descriptions that are provided for grammar, writing and coherence. The main part of the thesis will then be a practical analysis of discourse connectives in texts written by non-native speakers of English. The source for the practical part will be a collection of 161 formal application letters written by university students who are speakers of English as a second language and whose language proficiency corresponds to the level B2 in the CEFR. The thesis aims to define, analyze and classify the different kinds of connectives used by B2 speakers in detail. We will create a corpus of texts in which we will manually select and annotate individual connectives. The focus will be on the frequency of connectives, their realization forms (morphological and syntactic categories, e.g. word class), types (e.g. one-word/multi-word, primary/secondary connectives) and position in the sentence (initial/post-initial/medial/final). We will also include a syntactic characterization of discourse arguments (e.g. whether the connectives have an inter- on intra-sentential scope, whether the arguments are verbal or nominal), and a classification of the semantic types of relations that the connectives express (relations of expansion, contingency, comparison, temporal).

The reader of this paper will be familiarized with the methodology and results of the research in detail. The material will be analyzed quantitatively, and the chosen categories will be analyzed in detail using specific examples from the corpus of texts. The results will be

compared with our hypothesis (see 3.4) to see in which areas they correspond or differ. The aim of this thesis is to evaluate the use of discourse connectives as cohesive devices in formal texts written by non-native speakers of English, and to specify the level B2 in the Common European Framework of Reference in terms of written language. We will also aim to identify and evaluate the potential errors in the use of discourse connectives made by intermediate speakers of English.

2 THEORETICAL BACKGROUND

2.1 DISCOURSE

2.1.1 Approaches to Discourse, Discourse Analysis

Discourse analysis is recognized as one of the vastest but at the same time least defined topics in linguistics (Schiffrin, 1994: 1; Stubbs, 1983: 12). One of the reasons for this is that discourse has been defined across different academic disciplines. Apart from linguistics, discourse plays a role in fields such as sociology, anthropology, philosophy, social psychology, communication or artificial intelligence (Schiffrin, 1994: 1; Van Dijk, 2011: 1). The approaches to discourse and discourse analysis often overlap, but differ in aspects such as their origins, theoretical assumptions or methods for collecting and analyzing data.

According to Schiffrin (1994), there are two main approaches to the definition of discourse: the *formalist* (or structural) and *functionalist* approach. These two approaches are tied to two respective linguistic paradigms which have different assumptions about the nature of language and the way language should be studied. This is naturally reflected in the way discourse is defined (Leech, 1983: 46). The formalist approach considers discourse as a *level of structure higher than the sentence*, while the functionalist approach considers discourse as *language use* (Schiffrin, 1994: 20). Following is a more detailed explanation of the two theories and their approach to discourse and discourse analysis.

2.1.1.1 The Formalist Approach to Discourse

Formal linguists such as Chomsky (2006) regard language as a mental phenomenon, in which linguistic universals derive from a common linguistic inheritance and a built-in capacity to learn language. Formalists study language as an autonomous system. In the formalist approach, discourse is defined as “language above the sentence or above the clause” (Stubbs, 1983: 1). Discourse is basically a synonym to *text* and is comprised of lower *units* or *constituents* (clauses/sentences). These units function in relation to each other, not in relation to the context in which they appear. In other words, the processes external to the linguistic system are not considered to influence the internal organization of language.

Z. S. Harris, the first linguist to ever use the term *discourse analysis*, claimed that what distinguished discourse from a random sequence of sentences was precisely its structure, or the hierarchy of its smaller constituents (Harris, 1951 in Schiffrin, 1994: 24). These individual constituents must be related to each other within this hierarchy so that the principles of cohesion

and intelligibility are not violated. Discourse analysis then focuses on what these relations are and how they are expressed.

2.1.1.2 The Functionalist Approach to Discourse

Functionalists view language as a societal phenomenon, in which universals are derived from the use of language in society. Advocates of this approach, such as Halliday (1994), study language in relation to its social function. As opposed to the formalist approach, functionalists work with the assumption that there are processes external to the linguistic system that influence its internal organization. Discourse is analyzed in relation to the context in which it appears.

Discourse is defined as “language in use”, or “language in context” (Brown and Yule, 2012: 1). Functionalist analyses focus not on the formal properties of a language, but on what language is used for. Brown and Yule say that discourse analysis “cannot be restricted to the description of linguistic forms independent of the purposes or function which those forms are designed to serve in human affairs” (2012: 1). The functional approach has a much shorter tradition than the formalist approach. A unitary set of labels of the main functions of language has not yet been established. Since discourse is viewed as naturally occurring language, it is hard to draw borders between one function of language and another.

To sum up, there are two main approaches to discourse and discourse analysis – the formalist and functionalist approaches. In these approaches, discourse is respectively defined as either “language above the sentence” or “language in use”.

2.1.2 Text and Texture

The term *text*, which is by some linguists considered a synonym to *discourse*, has also been conceptualized in different ways. A *text* might be defined as the “verbal record of a communicative event” (Brown and Yule, 2012: 190), or as “any phenomenon that generates meaning through signifying practices” (Barker and Galasinski, 2001: 5). Fairclough (1995) defines text as a linguistic event where two fundamental processes are materialized: cognition and representation of the world (4). Due to the broadness of the term, linguists have been concerned with how speakers identify a text, or what the principle features of textuality are.

Halliday and Hasan (1976), the authors of one of the most comprehensive publications on the subject, *Cohesion in English*, use the term text to refer to “any passage, spoken or written, of whatever length, that does form a unified whole” (1). In this sense, any stretch of language that functions as a unity is a text – it can be anything from a single exclamation to a discussion between two interlocutors, or a full written book. A text is not a grammatical unit like the sentence or the clause, it is a semantic unit of language in use, defined by a unity of meaning,

not of form (1976: 2). If a text is defined by a unity of meaning, the question is how this unity of meaning is achieved. Halliday and Hasan (1976) write that speakers of a language have an instinctive ability to distinguish between a text and a random collection of unrelated sentences. There might, of course, be some degree of uncertainty,¹ but there seem to be some objective factors that define a text. De Beaugrande and Dressler (1981) talk about seven standards of texture: *cohesion*, *coherence*, *intentionality*, *acceptability*, *informativity*, *situationality* and *intertextuality*.² Out of these standards, the two features that have become widely accepted across authors are **cohesion** and **coherence** (Halliday and Hasan, 1976; Halliday, 1994; Schiffrin, 1994; Johnstone, 2008; Brown and Yule, 2012; Jaworski and Coupland, 2014, etc.). Since they are complex concepts, they will be explained in the following chapter.

In this thesis, discourse is regarded as a synonym to text, divisible into smaller segments that are connected by different kinds of relations.

2.1.3 Cohesion and Coherence

Although cohesion and coherence have both become key terms in discourse analysis, the conceptualization of the two terms and the relation between the two varies across publications and theories.

Cohesion is by some authors defined as the formal or surface relations in a text, i.e. the lexical, grammatical and syntactic structures that hold a text together. De Beaugrande and Dressler (1981: 1) define cohesion as “the way in which the components of the surface text, i.e. the actual words we hear or see, are mutually connected within a sequence”.

Coherence, on the other hand, is regarded as the deeper semantic relations of a text, or the way a text is made semantically meaningful. As Beaugrande and Dressler (1981: 4) write, coherence is concerned with “the way in which the components of the textual world, i.e. the configuration of concepts and relations which underlie the surface text, are mutually accessible and relevant”.

Jaworski and Coupland (2014: 155–157) also separate cohesion from coherence, arguing that cohesion is neither a sufficient nor a necessary condition for coherence. They define cohesion as “the inter-sentential semantic relations that link current items with preceding or following ones by lexical and structural means”. Since cohesion rests on grammatical forms

¹ A seeming non-relatedness of utterances might be unintentional, especially in speakers with a lower language proficiency.

² In other words, a text must be continuous, coherent and intelligible, it must be intended to be a text and accepted as such to be utilized in communicative interaction. A text is in some way informative, relevant to a current or recoverable situation of occurrence, and its production and reception depend on the participants' knowledge of other texts.

and dependencies, it is an invariant property of a text, independent of user and context. Coherence, on the other hand, is a cognitive category that largely depends on the user's interpretation of the text. While the term *cohesion* has been relatively well defined, *coherence* remains a vague concept (Jaworski and Coupland, 2014: 166). Sanders et al. (1993) hold that there is no satisfactory theory on coherence, because there is no unitary classification of coherence relations.

As James Robert Martin (in Schiffrin et al., 2001: 36) points out, early canonical work on cohesion moved beyond the structural resources of grammar. Martin provides a slightly different approach to the concepts of coherence and cohesion in relation to texture, seeing them as three interrelated concepts (cohesion – texture – coherence). Martin defines texture as the “process whereby meaning is channeled into a digestible current of discourse” (in Schiffrin et al., 2001: 35), and views cohesion as one aspect of it, alongside the text-forming resources of grammar and phonology. Texture is one aspect of the study of coherence, which he defines as the “process whereby a reading position is naturalized by texts for listener/reader” (in Schiffrin et al., 2001: 35). Apart from texture, the process of coherence involves the interlocutor's understanding and expectations about the social context that the text forms.

More recent research on the topic involves the work by Hoey (e.g. 1991), who proposes a lexically-based system for the analysis of cohesion. He argues that the lexical devices that are commonly associated with cohesion are all essentially types of *repetition*, and that a close relation in text is indicated when there is a high occurrence of lexical repetition across sentences. Apart from repetition, he mentions a list of other cohesive links such as *substitution*, *paraphrase*, *ellipsis* and *deixis*, saying that key sentences in a text can be identified by a higher number of these types of links in comparison to surrounding sentences. In this sense, the presence of lexico-grammatical features in a text (*cohesive links*) can be indicative of the surrounding context. The overtly-expressed features that provide for cohesion make the reader look into the surrounding context for a final interpretation of meaning.

Halliday and Hasan (1976), who are viewed as a canonical study of the branch of Systemic Functional Linguistics (SFL)³, and whose model will be adopted in this thesis, provide a semantic approach to the concept of cohesion and its relation to texture. Their approach to cohesion will be described in the chapter that follows (2.1.3.1).

³ The term cohesion is generally associated with two traditions in linguistics: Systemic Functional Linguistics (SFL) and Hartford-based stratification linguistics (Martin in Schiffrin et al., 2001: 35). Halliday and Hasan (1976) are viewed as a canonical study of the first branch and Gutwinski (1976) of the second.

2.1.3.1 Cohesion in Halliday and Hasan (1976)

In one of the fundamental works of English linguistics, *Cohesion in English* by Halliday and Hasan (1976: 4), cohesion is defined as a semantic concept that refers to “relations of meaning that exist within the text, and that define it as a text”. A relation of cohesion is set up when the interpretation of some element in the discourse is dependent on that of another, in a way that one cannot be effectively decoded without recourse to the other). Halliday and Hasan (1976) provide the following example to demonstrate the phenomenon:

(1) *Wash and core six cooking apples. Put them into a fireproof dish.* (ibid.: 3)

Here, the demonstrative pronoun *them* in the second sentence anaphorically refers to the noun phrase *six cooking apples* in the first sentence. The anaphoric function of the demonstrative pronoun gives the two sentences cohesion, so that the two sentences together constitute a text. Texture is provided by the cohesive relation which exists between the pronoun *them* and the noun phrase *six cooking apples*. Cohesion is achieved not by the demonstrative pronoun alone, but by the presence of both the referring item and the item it refers to (which are referred to as the *presupposing* and *presupposed* elements). One occurrence of a pair of cohesively related items is referred to as a **cohesive tie**. The presupposed element in a cohesive relation is most commonly a specific item in the immediately preceding sentence, but there are instances where one presupposing item may refer to several presupposed elements. Such a higher unit that is formed by two or more cohesive ties is called a **cohesive chain**.

Since the work by Halliday and Hasan (1976) is fundamental in English linguistics, their semantic concept of cohesion will be adopted in this thesis. Coherence will be considered as an umbrella term for both the grammatical and semantic unity of discourse (text).

2.1.4 Lexical and Grammatical Cohesion

Halliday and Hasan (1976) describe five general cohesive devices or strategies that speakers use to show how the meanings of parts of different sentences are related to each other. They divide these strategies into grammatical and lexical groups, according to the way they express cohesive ties. The phenomena that fall under grammatical cohesion are: **reference**, **substitution**, **ellipsis**, and **conjunction**; **reiteration** and **collocation** are grouped under lexical cohesion.

In Halliday and Hasan’s terminology, conjunction refers to the category of connectives, and since they are the focus of this thesis, these are discussed in a separate chapter (see 2.1.4.3). Following is a brief overview of the other types of cohesion mentioned by Halliday and Hasan.

2.1.4.1 Lexical Cohesion: Reiteration, Collocation, General Nouns

Halliday and Hasan mention two basic means of lexical cohesion: reiteration and collocation. **Reiteration** is a form of lexical cohesion that can involve the direct repetition of a lexical item on the one hand, the use of a general word to refer back to another lexical item on the other, or the use of a synonym, near-synonym, or superordinate term, which are scalar phenomena. In this sense, reiteration is related to reference, but it is considered a lexical means of cohesion, that is, a cohesive effect achieved by the selection of vocabulary, and not of grammar.

Collocation is cohesion that is achieved through the association of lexical items that regularly co-occur (Halliday and Hasan, 1976: 284). The cohesive effect of collocates is mostly due to their tendency to share the same lexical environment. Lexical items that tend to appear in similar contexts often occur in adjacent sentences and thus generate a cohesive force.

Halliday and Hasan also include the class of **general nouns** as a device on the borderline between grammatical and lexical cohesion (1976: 274). This class is a small set of nouns that have a generalized reference within a major noun class, such as *child*, *boy*, *girl* for [human], *move* for [action], or *question*, *idea* for [fact]. These nouns may have a cohesive function since they can refer to the same referent as the item which they presuppose. General nouns in this function are almost always accompanied by the reference item *the*. The construction ‘*the* + general noun’ functions together like an anaphoric reference item. Consider the following example:

(2) *I've been to see my great-aunt. The poor girl's getting very forgetful these days.*

(Halliday and Hasan, 1976: 274)

As can be seen, the noun phrase with the construction ‘*the* + general noun’, *the poor girl*, anaphorically refers to the noun phrase *my great-aunt* in the first sentence. Thus, this strategy combines the grammatical means of the definite article and the lexical means of repetition using a general noun to form a cohesive tie between sentences.

2.1.4.2 Grammatical Cohesion: Substitution, Ellipsis, Reference

Halliday and Hasan (1976) adopt a straightforward model of co-reference: **substitution** refers to the replacement of one expression for another in a text. **Ellipsis**, which is very closely tied to substitution, is the omission of an item. The authors write that the two are essentially the same process, since ellipsis can be interpreted as a form of substitution by zero (Halliday and Hasan, 1976: 88). Both substitution and ellipsis can be nominal, verbal, or clausal. The mechanisms of these phenomena, especially ellipsis, are rather complex, and would require a much larger study to be explained in detail.

An example of **reference** has already been provided above (1):

Wash and core six cooking apples. Put them into a fireproof dish. (Halliday and Hasan, 1976: 3)

In this example, the demonstrative pronoun *them* in the second sentence refers anaphorically to the noun phrase *six cooking apples* in the first sentence, creating a cohesive tie between the two sentences, which together constitute a text. Halliday and Hasan further distinguish between endophoric (textual) and exophoric (situational) reference, the former of which can be either anaphoric (referring to preceding text), or cataphoric (referring to following text). They mention three basic types of reference: *personal*, *demonstrative*, and *comparative*. Personal reference is reference through the category of person, demonstrative reference is reference by means of location on a scale of proximity, and comparative reference is indirect reference by means of identity or similarity (1976: 37).

2.1.4.3 Grammatical Cohesion: Conjunction

Halliday and Hasan do not speak directly of discourse markers or discourse connectives, but their analysis of the cohesive device of conjunction includes words that serve as discourse connectives, such as *and*, *but*, *because*, *to sum up*.⁴ It can be said that Halliday and Hasan's category of conjunction is analogous to the class of discourse connectives.

Reference, substitution and ellipsis all signal words that are already available in text, whether they form a relation between linguistic forms or meanings. Conjunction, however, is a much more complex type of semantic relation. As Halliday and Hasan write, it is a "specification of the way in which what is to follow is systematically connected to what has gone before" (1976: 227). Whereas the grammatical cohesive features mentioned above (reference, repetition, substitution, ellipsis) express relations through anaphoric or cataphoric ties, conjunctive items are not cohesive in themselves, they "express certain meanings which presuppose the presence of other components in discourse" (Halliday and Hasan, 1976: 236). Conjunctive relations are not tied to any sequence in an expression; the conjunctive relation between two sentences is not dependent on the order of these sentences. For example, one relative expression of time may have several different realizations. The sentence that refers to the event earlier in time might, in fact, come after the sentence that refers to the later event:

(3) ***After** they had fought the battle, it snowed.*

(4) *They fought a battle. **Afterwards**, it snowed.* (Halliday and Hasan, 1976: 227)

⁴ More examples of conjunctive elements are provided in 2.2.3.

Though one relation may be expressed by a variety of synonymous structures, the logic of the connection is maintained, which means that it is not the conjunctive expression itself, but the underlying semantic relation that carries the cohesive power.

2.1.4.4 Cohesion and Coherence in Evaluating Writing

Cohesion and coherence have become a subject of great interest in terms of teaching and evaluating writing. For example, Witte and Faigley (1981) investigated the relation between patterns of cohesiveness and quality of writing. Their results showed that in comparison to lower-rated college essays, high-rated essays were denser in cohesive ties. Neuner (1987) analyzed cohesive ties and chains in essays by college freshman students. The findings of his research imply that even though cohesive ties are not the distinguishing features of good or bad writing, there are several features of cohesive chains that characterize good writing: the presence of longer chains, a greater variety of words, and greater maturity of word choice.

2.1.5 Defining the Basic Unit of Discourse

The identification of the basic unit of discourse varies across authors. Some authors (Halliday and Hasan, 1976) divide discourse into smaller units called sentences, others (Fraser, 1996) into utterances. Barbara Johnstone (2002) provides an overview of some of the approaches to “chunking” discourse into units. The approaches highlight different aspects of discourse and are associated with different theories of discourse structure. An overview of some of the approaches is in the chart below.

Basic Unit of Discourse	Definition
Sentence	The expression of a complete thought, containing a subject and a predicate
Clause	The construction subject + finite verb + verbal complements
“t-unit” (Hunt, 1966)	The construction main clause + embedded or dependent clauses
Intonation unit	A set of words that expresses a single cognitive ‘focus of consciousness’; surrounded by pause
Utterance	A stretch of speech with a single intonation contour, bonded by pauses and/or constituting a single unit of meaning
Tone group	A stretch of speech consisting of one prominent pitch nucleus and one or more other syllables

Table 1: Basic Units of Discourse According to Johnstone (2002)

The approaches use different criteria in their definition of the basic unit of discourse. The units *sentence*, *clause* and *t-unit* are based on structural features, the *intonation unit* and *tone group* on acoustic features (and are used mostly in reference to oral discourse); the term *utterance* is

based on semantic features of discourse. The segments of text into which authors “chunk” discourse vary in length – some authors even consider the paragraph as a possible unit of discourse. As can be seen, the only consensus is that discourse is divisible into smaller units or segments, but what the basic unit is has not yet been agreed on.

2.1.5.1 The Penn Discourse Treebank and Discourse Arguments

The Penn Discourse Treebank (Prasad et al., 2007) is a large corpus annotated with information related to discourse structure and semantics, focused on encoding discourse relations. It is built on top of the Penn Treebank (Marcus et al., 1993) and the PropBank (Palmer et al., 2005), which use semantic and syntactic features to evaluate significant algorithms in discourse. The PDTB (Prasad et al., 2007) draws on research such as Mann and Thompson (1988), Polanyi and van den Berg (1996), Asher and Lascarides (1998), Webber and Joshi (1998), and Webber et al. (2003). The PDTB (Prasad et al., 2007) contains annotations not only of explicit discourse relations, but also implicit relations across the entire corpus.

In the PDTB (Prasad et al., 2007), the basic unit of discourse is called a *discourse argument*. Discourse arguments are defined using the *minimality principle*. According to this principle, a discourse argument includes only the amount of information (i.e. the number of clauses or sentences) that is *minimally required* and at the same time *sufficient* to complete the semantics of the relation. In other words, a discourse argument is the shortest span required for the interpretation of a discourse relation. Any other span of text that is perceived to be relevant, but not necessary, is called *supplementary information* (Prasad et al. 2007: 14). There are also no restrictions as to where one finds the first argument of a relation. It can be adjacent or non-adjacent to the second argument (i.e. the argument that hosts the connective), it can be continuous or discontinuous (Benz et al., 2012: 190). The PDTB (Prasad et al., 2007) base their concept of discourse relations on Asher (1993), who defines discourse arguments from a semantic perspective as *abstract objects*. Abstract objects can express states, events, situations, propositions, etc., and are syntactically realized by clauses or phrases. Asher’s (1993) hierarchy of abstract objects is shown below.

Abstract Objects								
Eventualities		Fact-like Objects			Proposition-like Objects			
Events	States	Situations	Facts	Possibilities	Desires	Propositions	Commands	Questions

Table 2: Hierarchy of Abstract Objects According to Asher (1993: 57)

2.1.5.2 Syntactic Features of Discourse Arguments

Some authors define connectives according to the syntactic features of the arguments they join. In the strictest sense, discourse connectives are defined as linguistic expressions that join two verbal arguments, i.e. segments of text that contain a finite verb (Rysová, 2015: 64). This is what distinguishes the connective function of an expression (*He is handsome **and** he knows how to cook*) from its structural function on the phrasal level (*He is tall **and** handsome*).⁵ However, when taken from a semantic perspective, discourse arguments do not always have to be verbal. Consider the following examples:

(5) *I didn't go to school today **because** I felt ill.*

(6) *I didn't go to school today *because of* my illness.*

In example (5), the subordinating conjunction *because* joins two verbal arguments (Arg1: *I didn't go to school today*; Arg2: *I felt ill*). In example (6), the second argument is expressed by a nominalized phrase (*my illness*), which is possible thanks to the addition of the preposition *of* to the connective (*because of*). Although the second sentence contains only one finite verb form (*didn't go*), the meaning of the utterance and the causal relation between the two arguments is maintained.

Although discourse arguments are in most cases realized by clauses containing a finite verb form, some connectives allow for the nominalization of one of the arguments. The difference between verbal and nominal arguments is a syntactic, not a semantic one. In terms of discourse relations, both options are semantically equal.

In this thesis, we adopt the model of the PDTB (Prasad et al., 2007) and consider the *discourse argument* as the basic unit of discourse. Discourse arguments will be defined as abstract objects that express complete thoughts or events, facts, propositions, etc., and are realized by finite verb clauses (*verbal arguments*) or nominalized phrases (*nominal arguments*).

2.2 DISCOURSE RELATIONS

Discourse connectives indicate that two units of discourse stand in a particular coherence relation (Unger, 1996: 410). The question is whether discourse connectives create meaning or simply indicate existing relations. Sanders et al. (1993) argue that there is no one-to-one correspondence between the meaning of particular discourse connectives and particular coherence relations. However, the meaning of the connective and the coherence relations which

⁵ For the distinction between the structural and connective use of an expression, see 2.3.2.1.

hold between the two discourse units it links must be compatible. Sanders et al. (1993) formulate the point in the following way:

... connectives certainly play a role in guiding the interpretation of the relation: the coherence relation that is assumed between the segments must be compatible with the meaning of the connective and with the meaning of the segments. (ibid.: 94)

In other words, discourse connectives do not create meaning, but they put constraints on the range of coherence relations that the hearer infers between the two segments involved (Unger, 1996: 410). Discourse relations exist even if they are not expressed by a discourse connective, which will be explained in the following chapter.

2.2.1 Explicit and Implicit Relations

Coherence relations may be overtly expressed in text, but they can also be inferred from the context. The Penn Discourse Treebank (Prasad et al., 2007) distinguishes between *explicit* and *implicit* relations, which are annotated as explicit and implicit connectives. Explicit relations are those which are grammatically expressed in discourse. Implicit relations, on the other hand, are not realized by overt linguistic means. If the relation is implicit, the meaning of the text connections must be deduced by the reader. Consider the difference between the two following examples:

(7) *I don't like him. He stole my glasses once.*

(8) *I don't like him **because** he stole my glasses once.*

In both cases, the semantic relation between the arguments is that of cause-and-effect. The difference is in the way this relation is expressed. In (8), the relation is overtly expressed using a causal discourse connective, *because*. In (7), there is no such overt marking, so the semantics of the two sentences must be logically inferred by the reader using other cohesive means (e.g. personal reference, *him–he*). Although the meaning of the utterance is simple, the semantic meaning of the clause with the explicit connective is more easily recoverable. Generally, explicit means of discourse relations are easier to process than implicit means. Irwin (1980) has carried out various studies on discourse processing with the variables of implicit and explicit connectives, focused specifically on children and younger adults. The results of her studies indicated a higher level of comprehension when explicit relations were present, as compared to implicit test groups.

2.2.2 External and Internal Relations

Halliday and Hasan (1976) further make a distinction between external and internal relations. External relations are defined as those relations that are inherent in the phenomena of language, while internal as those relations that are inherent in the communication process. In relation to the functional components of meaning, external meaning is roughly analogous to referential semantic meaning, while internal meaning is nonreferential pragmatic meaning (240–241). Consider the following sentences:

(9) *She was never really happy there. So she's leaving.*

(10) *She'll be better off in a new place. – So she's leaving?* (Halliday and Hasan, 1976: 241)

In example (9) above, there is a causal relation between the two events, with the meaning *because she was not happy → she is leaving*. In the example (10), there is also a causal relation, but only within the communication process, with the meaning *you refer to her being better off in a new place → I conclude that she is leaving*. The two categories of external and internal connectives are stable, independent of the type of coherence relation.

2.2.3 Semantic and Pragmatic Types of Relations

Halliday and Hasan (1976: 250) classify connectives into four main types according to the kind of relation they express in text: **additive**, **adversative**, **causal**, and **temporal**. **Additive** relations are derivable from the relation of coordination. **Adversative** relations are relations with the meaning 'contrary to expectation' (Halliday and Hasan, 1976: 250). This expectation may be derived either directly from the content of what is being said, or from the communication process or situation. **Causal** relations include the semantic relations of result, reason and purpose. **Temporal** relations may express both external and internal sequences of events in time, as well as sequences of the communication process. Below is a table of these categories and some of their subcategories, including examples of connectives by which they can be expressed.

Relation	Subcategories	Examples of Connectives
Additive	Additive simple	<i>And, nor, or else</i>
	Emphatic	<i>Further, furthermore</i>
	Comparison	Similarity: <i>likewise, similarly</i> Dissimilarity: <i>on the other hand, by contrast</i>
	Apposition	Expository: <i>that is, I mean</i>
		Exemplificatory: <i>for instance, thus</i>
Adversative	Adversative ‘proper’	<i>Yet, though, however, nevertheless</i>
	Contrastive	<i>In fact, as a matter of fact, on the other hand</i>
	Correction	Of meaning: <i>instead, rather</i> Of wording: <i>at least, rather</i>
	Dismissal	<i>In any case, anyhow</i>
Causal	General	<i>So, then, hence, therefore, consequently</i>
	Result	<i>As a result, in consequence</i>
	Reason	<i>For this reason, on account of this</i>
	Purpose	<i>For this purpose, with this in mind</i>
Temporal	Sequential	<i>Then, next, after that</i>
	Simultaneous	<i>Just then, at the same time</i>
	Preceding	<i>Previously, before that</i>

Table 3: Table of Semantic Relations According to Halliday and Hasan (1976: 242)

As can be seen, each of the cohesive meanings can be conveyed through a variety of linking words, e.g. a causal relation can be expressed by the connectives *so, thus, hence, therefore*, etc. There is multiplicity not only in the direction *function* \rightarrow *form*, but also in the direction *form* \rightarrow *function*. One linguistic form of a connective may express several different kinds of semantic relations (e.g. the expression *on the other hand* may be additive or adversative; the connective *then* can express causal or temporal relations, the connective *when* can express temporal, causal, and conditional relations). A one-to-one assignment of semantic roles of connectives is, therefore, not possible.

The PDTB (Prasad et al., 2007) offers a slightly different categorization of semantic relations, which are called *senses*. The PDTB divides senses into four groups: **temporal**, **contingency**, **comparison** and **expansion**, each of which has their own subcategories. Below is a diagram with an overview of the main categories and subcategories of senses according to the PDTB.

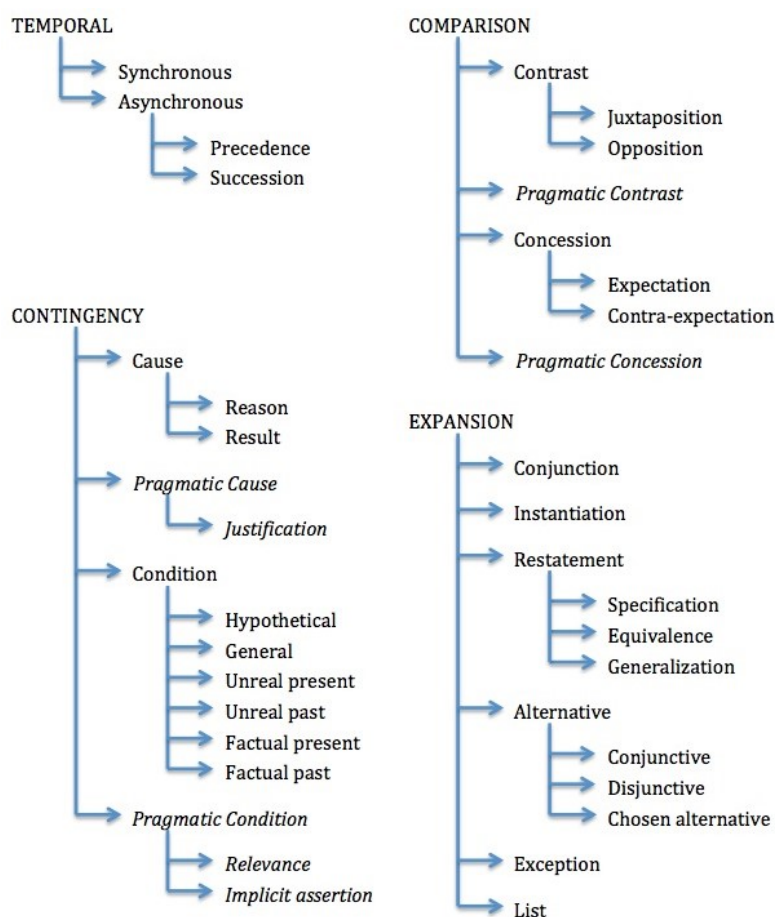


Figure 1: Semantic Relations According to the PDTB (Prasad et al., 2007)

The category of **temporal** relations is shared by the PDTB (Prasad et al., 2007) and Halliday and Hasan (1976). The PDTB further divides temporal relations according to whether the events are temporally ordered or overlapping into *asynchronous* (temporally ordered) and *synchronous* (temporally overlapping). Asynchronous relations include connectives that express e.g. precedence (*before*) or succession (*after*). Synchronous relations apply when the connective indicates that the situations overlap and does not specify the form of overlap (whether the situations started and ended at the same time, or whether one was embedded in the other, etc.). Among typical synchronous connectives are *while* and *when*.

The class of **contingency** applies when the connective indicates that one situation influences the other causally, and so it loosely corresponds to the category of causal relations in the classification by Halliday and Hasan (1976). Relations of contingency include the subtypes of *cause* (plus pragmatic cause) and *condition* (plus pragmatic condition).

Connectives of **comparison** indicate a discourse relation that is established between two arguments to highlight the salient differences between them. Semantically, the two arguments are independent of the connective or established relation. The two main subtypes of comparison are *contrast* and *concession*. The relation of contrast is set when the two arguments share a

property and the difference between them is highlighted with respect to it (e.g. *but, whereas*). The relation of concession occurs when the differences raised by one argument are denied by the other (e.g. *although, even though*).

The last group of relations according to the PDTB is **expansion**. The class of expansion covers relations that expand the discourse, i.e. move it forward. Expansion has the most subtypes, including: *conjunction, instantiation, restatement, alternative, exception, and list*. The term *conjunction* (cf. *conjunction* as defined by Halliday and Hasan, 1976) is used when the second argument provides additional and new information that is related to the first argument, but not by any of the other subtypes of expansion. The prototypical connectives of this class are *also, in addition, additionally, further*.

Connectives can also be used to relate the use of the arguments to one another or the use of one argument with the sense of the other (Prasad et al., 2007: 27). For these pragmatic uses of connectives, the PDTB defines the **pragmatic senses of cause, condition, contrast and concession**. In these pragmatic uses of connectives, there is no direct causal influence between two situations. Consider the following example of pragmatic condition. The connective is in **bold**, the first argument (arg1) is underlined, and the second argument (arg2) is in *italics*.

(11) ***If you are thirsty***, there's beer in the fridge. (Prasad et al., 2007: 32)

Although the connective *if* normally introduces a condition, there is no causal relation between the two arguments above (Arg1: you are thirsty → Arg2: there's beer in the fridge). Argument 1 holds independently of argument 2 and the relation between the two is only pragmatic.

2.3 DISCOURSE CONNECTIVES

2.3.1 Terminology

The expressions that we call *discourse connectives* have been studied in many languages and across a variety of genres and contexts. In the broadest sense, discourse connectives are defined as the explicit means of signaling semantic relations in a text (Rysová, 2015: 54). However, individual scholars differ in their definition of discourse connectives, depending on their approaches to discourse and methods of analysis. One way to approach discourse analysis is to proceed from concrete linguistic analysis to the generalization of the structure of discourse. Authors such as Prasad (2008, 2010) and Stede and Neumann (2014) base their analysis on the identification of linguistic means that have the function of discourse connectives. For such approaches, a definition of discourse connectives is crucial.

The fact that discourse connectives are not easily defined becomes apparent in the problematics of the denomination itself – the term *discourse connectives* is not stable. These

expressions appear in international literature under different names, including: *discourse connectives* (Blakemore, 1987, 1994; Rouchota, 1996; Celle and Huart, 2007), *discourse connectors* (CEFR, 2001; Granger and Tyson, 1996; Pečený, 2017), *discourse particles* (Fischer, 2006), *discourse markers* (Blakemore, 2004; Schiffrin, 1987; Urgelles-Coll, 2010), *discourse operators* (Redeker, 1991), *pragmatic markers* (Fraser, 1988, 1990, 1996, 1998), *conjunctive adverbials* (Chen, 2006), *linking adverbials* (Shaw, 2009), *pragmatic operators* (Ariel, 1994), *pragmatic connectives* (Van Dijk, 1979), *phatic connectives* (Bazanella, 1990), *semantic conjuncts* (Quirk et al., 1985), *conjunctive elements* (Halliday and Hasan, 1976), to name just a few. It is not only the terminology that differs across authors, but the aspects from which connectives are defined, and the approach to their role in discourse. The following chapter is devoted to explaining some of the main approaches to discourse connectives in international and English literature, and it aims to conclude with a definition that is used for the analytical part of this thesis.

2.3.2 Approaches to Discourse Connectives

Schiffrin (2001) mentions three influential perspectives of discourse analysis which differ in their approach to discourse connectives: Halliday and Hasan's semantic approach (1976), Fraser's pragmatic approach (1990, 1998), and Schiffrin's own sociolinguistic approach (1987).

2.3.2.1 Halliday and Hasan's (1976) Approach to Discourse Connectives

As explained previously, Halliday and Hasan (1976) take a semantic perspective on discourse (see 2.1.3.1, 2.1.4). They refer to the category as *conjunction*, but their analysis includes linguistic expressions that have since then been called discourse connectives. They propose a set of cohesive devices (*reference, repetition, substitution, ellipsis, conjunction*) that help to form a text by indicating semantic relations in an underlying structure of ideas. Halliday and Hasan (1976) draw a sharp distinction between the sentence and the text, which is reflected in a strict separation between the structural role of a word and the cohesive role of a word. Consider the following sentences:

(12) *He went to buy apples **and** pears.*

(13) *'I wonder if all the things move along with us?' thought poor puzzled Alice. **And** the Queen seemed to guess her thoughts, for she cried 'Faster! Don't try to talk!'*
(Halliday and Hasan, 1976: 235)

Example (12) is an instance of structural coordination at a sentential level. Here the conjunction *and* coordinates two nouns, whose order could be easily swapped without changing

the meaning of the utterance (*He went to buy pears **and** apples*). The pair of items joined by coordination (whether they are nouns, noun phrases, verbs, clauses...) functions as a single complex element of structure. If more elements were added to the list (*He went to buy apples **and** pears **and** milk **and** bread*), they would still constitute a single element in the structure of the clause – in this case, the object. In example (13), on the other hand, the sentences cannot be rearranged without losing cohesion. They follow one another as the text unfolds, and if another sentence followed, it would be link up to the second sentence; the three would not form a whole. Here, the role of *and* is cohesive – it expresses a dependency between two propositions, creating texture. A word in its cohesive role typically occupies the first position in the sentence, and its meaning extends over the entire sentence (Halliday and Hasan, 1976: 232).

2.3.2.2 Fraser's (1990, 1998) Approach to Discourse Connectives

Fraser's (1990, 1998) approach to discourse connectives is a pragmatic one (hence the term *pragmatic markers*), centered around the question of how a pragmatic marker can relate the message conveyed by a sentence to the message of a sentence prior to it. He provides a classification of the types of pragmatic meanings and describes how some markers dictate an interpretation of the message conveyed (Fraser, 1998: 302). His framework presupposes a strict separation between content meaning (referential) and pragmatic meaning (the speaker's communicative intention). Pragmatic meaning can be expressed by three different sets of pragmatic markers: basic pragmatic markers, commentary pragmatic markers, and parallel pragmatic markers. Discourse connectives are a type of commentary pragmatic markers, defined as "a class of expressions, each of which signals how the speakers intends the basic message that follows to relate to the prior discourse" (1990: 387).

2.3.2.3 Schiffrin's (1987) Approach to Discourse Connectives

Schiffrin (1987) takes a more sociolinguistic perspective on the analysis of discourse connectives. Her methodology aims at providing an account for the use and distribution of forms in discourse. She views discourse not only as a unit of language, but as a process of social interaction (Schiffrin, 2001: 56). In her analysis, discourse markers are defined as a set of "sequentially dependent elements that bracket units or talk, i.e. non-obligatory utterance-initial items that function in relation to ongoing talk and text" (Schiffrin, 1987: 31). These include conjunctions (e.g. *and*, *but*, *or*), interjections (e.g. *oh*), adverbs (e.g. *now*, *then*), but also lexicalized phrases (e.g. *y'know*, *I mean*).

2.3.3 Aspects of Definition

2.3.3.1 Word Class

Connectives as such are not a grammatical class and can be realized by single words, multi-word expressions, and even clauses or sentences. Most authors agree that connectives can be expressed by coordinating and subordinating conjunctions (e.g. *and*, *but*, *or*; *so*, *since*, *because*, *while*), prepositional phrases (*as a consequence*, *in particular*, *after all*, *on the other hand*), and adverbs (*then*, *afterwards*, *before*) (Zikánová et al., 2015; Halliday and Hasan, 1976; Prasad et al., 2008, 2010; Fraser, 1999). Some authors include more peripheral classes into the definition, such as particles and lexicalized phrases (Schiffrin, 1987) or idioms, such as *still and all* or *all things considered* (Fraser, 1999: 943). The Prague approach (Zikánová et al., 2015) even mentions elements formed by letters or numbers (*a*), *b*), *1.*, *2.*, *3.*) and punctuation marks (*the colon [:]* and *the dash [–]*) (16). Fraser (2009), on the other hand, excludes syntactic structures and prosodic features such as stress, pauses, intonation, or non-verbal expressions, from the definition. The difficulty arises in more complex expressions that cannot be easily categorized by word class, which is why authors usually do not restrict themselves to the grammatical or syntactic features of connectives, and consider other aspects of definition, such as their semantic or pragmatic meaning.

2.3.3.2 Pragmatic Aspect

Some authors define connectives according to whether they carry conceptual or procedural meaning (see 2.3.4). Expressions with a conceptual meaning specify a defining set of semantic features, as for example the nouns *boy* or *hypothesis*. Expressions with a procedural meaning specify how the segment they introduce is to be interpreted relative to the prior segment (Fraser, 1999: 944). Authors such as Fraser (1996, 1999), Rouchota (1996), or Blakemore (1987) agree that connectives are not part of the propositional content of the message conveyed, and that they have a core procedural, not conceptual meaning. As Fraser (1999) writes, “when an expression functions as a discourse marker it relates two discourse segments and does not contribute to the propositional meaning of either segment” (944). Consider the connectives in the following sequences:

- (14) *I want to go to the movies tonight. **After all**, it's my birthday.*
- (15) *John will try to come on time. **All the same**, he is going to be reprimanded.*
- (16) *Harry is quite tall. **In contrast**, George is quite short. (ibid.: 944)*

In these sequences, the connectives in bold could be omitted without a change in the propositional content of the segments. They serve as lexical clues that indicate a coherence relation. Their more specific interpretation is 'negotiated' by the linguistic and conceptual context. This means that the one linguistic form (e.g. *but*) does not represent several separate connectives, but merely different interpretations which are formed by the context.

Urgelles-Coll (2010) or Blakemore in her later work (2002) argue that some connectives can in fact have a conceptual meaning: “there are expressions which encode procedures, but which contribute to what is thought of as truth conditional content” (Blakemore, 2002: 4). However, Blakemore’s newer research views connectives from a broader perspective, including more complex peripheral expressions such as *in other words* or *as a result*, which will not be regarded as connectives here. In this thesis, the core meaning of connectives is considered procedural.

2.3.3.3 Other Aspects

Due to the complexity of connectives, authors tend to define them from several different aspects. Redeker (1991) defines a connective as a “word or phrase that is uttered with the primary function of bringing to the listener's attention a particular kind of linkage of the upcoming utterance with the immediate discourse context” (1168). She further specifies that a connective is a unit, usually clausal, which is intonationally and structurally bounded. To define connectives, Redeker (1991) provides a list of expressions that are not connectives:

- clausal indicators of discourse structure (e.g. *let me tell you a story, as I said before, since this is so*),
- deictic expressions⁶ – as far as they are not used anaphorically (e.g. *now, here, today*)
- anaphoric pronouns and noun phrases,
- any expressions whose scope does not exhaust the utterance (ibid.: 1168).

Fraser (1999) classifies connectives in terms of their syntactic, semantic and pragmatic properties. He defines connectives as expressions drawn from the syntactic classes of conjunctions, adverbials, or prepositional phrases, which have the syntactic properties associated with their class membership, a procedural meaning, and co-occurrence restrictions which are in complementary distribution with their conceptual counterparts (Fraser, 1999: 946). He further specifies that connectives are free morphemes that signal a specific message either

⁶ The boundary between the adverbial and connective uses of some expressions (e.g. *now, then*) is fuzzy. Cf. *I'll do it **now**.* / *It's nice here. **Now** our street isn't that nice.* (Redeker, 1991: 1146).

about or in addition to the basic message conveyed. Fraser (2009: 6) specifies several conditions to be a connective:

- a) A connective is a lexical expression.
 - This excludes syntactic and prosodic features such as stress, pauses, intonation or non-verbal expressions.
- b) In a sequence of discourse segments, the connective occurs as part of the second segment.
 - Fraser does not restrict connectives to initial position though, they may occur medially/finally within the second segment of a sequence of discourse.
- c) Connectives do not contribute to the semantic meaning of the segment.
 - In contrast to other pragmatic markers (such as *I promise, frankly, allegedly*), connectives are not part of the propositional meaning of an utterance.

These conditions lead to the exclusion of sentence adverbs (*certainly, surely, definitely*), modal particles (*indeed*), focus particles (*just, even, only*), evidential adverbs (*reportedly, according to*), and attitudinal adverbs (*frankly, stupidly, cleverly*), which all do not represent a semantic relation between adjacent segments. Further, complementizers (e.g. *that – I believe that John is right*) and topic orientation markers (relations that specify relations between the discourse such as *incidentally, that reminds me*) are also not considered connectives, since they express a grammatical relation, not a semantic one (Fraser, 2009: 7). Similarly, vocatives signal a message in addition to the primary message conveyed, and not a relation between segments. Interjections (*damn, hey, wow*) are also excluded, since they constitute an entire, separate message (Fraser, 1999: 942).

Urgelles-Coll (2009) defines connectives from phonological, syntactic, semantic, and pragmatic aspects:

- Phonologically, they are short and reduced.
- Syntactically, they are not integrated and can be omitted without affecting the grammaticality of the sentence.
- Semantically, they do not usually affect the truth conditions of the propositions they appear in.

2.3.4 Coherence and Relevance Theory

In his work *Discourse Connectives: What do they link?* (1996), Rouchota provides an overview of two basic theories on how connectives contribute to the interpretation of an utterance: **relevance theory** and **coherence theory**. The main question Rouchota poses is

whether connectives link an utterance to previous discourse units, or whether they link an utterance to a context (which does not necessarily have to be verbal). The fact that the two theories take different views on communication has consequences for the definition of connectives suggested by each framework.

2.3.4.1 Coherence-based Approach

The most important property of texts according to coherence theorists is that the sections of a text are linked together by means of different kinds of relations, such as cause, sequence, condition, evidence, or elaboration (Rouchota, 1996: 1). The process of understanding an utterance stems from recovering the coherence relations that hold between individual discourse units. Consider the following example:

(7) *I don't like him. He stole my glasses once.*

The hearer must interpret this as a causal relation with the following meaning: since the person in question stole the speaker's glasses once, the speaker doesn't like him. For coherence theorists, the role of connectives is to make implicit coherence relations explicit. For example, a simple addition of the connective *because* would make the relation between the discourse units of the previous utterance easily recoverable:

(8) *I don't like him **because** he stole my glasses once.*

Among advocates of this approach are Sanders, Spooren and Nordman (1993), Mann and Thompson (1988), or Fraser (1990).

2.3.4.2 Relevance-based Approach

Christoph Unger (1996: 413) argues that the coherence-based approach to discourse cannot be an adequate framework for the analysis of discourse connectives, since coherence relations are not directly part of the structure of discourse. Instead, the perception of coherence relations results from the way relevance is optimized over discourse.

Within the relevance-theoretic account, there are two issues that play a role in the interpretation of utterances. The first is related to the selection and the role of context. A speaker usually has a specific interpretation an utterance in mind and expects the reader to arrive at the same interpretation. For this to happen, the reader must process the utterance in the right context. The speaker may assume that the hearer will choose the appropriate context on his own, but he may also direct the hearer towards the right interpretation by providing a set of immediately accessible assumptions. Connectives are some of the linguistic devices that the speaker may use for this purpose.

The second issue has to do with the processing effort needed to interpret an utterance. As mentioned previously, Blakemore (1987, see 2.3.3.2) claims that connectives encode procedural rather than conceptual information, and that the function of connectives is to guide the process of interpretation by specifying certain properties of the context:

Cohesive devices do not themselves create meaning; they are clues used by speakers and hearers to find the meaning which underlie surface utterances. (Blakemore, 1987: 9)

In other words, connectives encode procedural information to the hearer and lead him in a certain direction in his search for the optimally relevant interpretation of an utterance in which the connective occurs.

2.3.4.3 Connectives in Coherence and Relevance Theory

To sum up the approaches, the coherence theory view connectives as linking elements of units of discourse (which are usually understood as two consecutive clauses) that indicate specific coherence relations. In relevance theory, connectives are viewed as linking units of discourse and context, and they encode procedural meaning. What is interesting to note is that there is a similarity in the way the two approaches analyze the role of connectives in interpreting utterances. Both accounts view connectives as having a constraining function. In the coherence-theoretical approach, connectives can be used to constrain the propositions that express relations of coherence that the hearer must recover to interpret an utterance. In the relevance-based approach, connectives are used to constrain the process of interpretation by guiding the hearer towards the intended interpretation or context.

Here connectives are defined as one-word or multiword linking units that signal semantic relations in text, have a pragmatic constraining function and carry procedural meaning.

2.3.5 Types of Discourse Connectives

Although many authors have attempted to define and classify connectives into groups according to their common semantic, syntactic or grammatical features, there is no singular classification of their types. Some consensus exists regarding the types of semantic relations that connectives express, but since the inventory of relations is so vast, it is hard to form a limited number of separate and clear-cut categories. As a continuation, some of the classifications will be listed and explained.

2.3.5.1 Primary and Secondary Connectives

In the Prague annotation scenario, Rysová (2015: 73) distinguishes between *primary* and *secondary* connectives. **Primary connectives**, which are the core category, are mostly one-word, grammaticalized expressions that are morphologically inflexible, lexically stable, and do not normally act as grammatical constituents of a sentence. They belong to the word classes of coordinating and subordinating conjunctions (e.g. *and*, *or*, *because*), adverbs (e.g. *however*, *moreover*, *therefore*) and prepositional phrases (e.g. *to sum up*, *in conclusion*, *on the other hand*). Similarly to sentence modality markers, they are ‘above’ or ‘outside’ the proposition. Primary connectives are often considered as the prototypical category of discourse connectives, since they are used much more frequently in expressing semantic relations in discourse than secondary connectives.

Secondary connectives are mostly multiword, non-grammaticalized phrases (cf. the primary connective *because* vs. the secondary connective *the reason is...*). Like primary connectives, they have a linking function in discourse, but are much more complex in definition and scope, and therefore are considered as the peripheral category of connectives. The class of secondary discourse connectives is a very heterogeneous one, containing expressions that function as clause elements (e.g. *because of this*), sentence modifiers (*simply speaking*), or even whole sentences (e.g. *The condition is clear*) (Rysová and Rysová, 2014). As Rysová (2015: 73) points out, secondary connectives function as whole structures, not as individual words in text. In the abovementioned example of a secondary connective, *the reason is*, the noun *reason* is not a connective by itself, but only functions as one in relation to the multiword whole.

Rysová (2015), who provides a detailed study of secondary connectives in Czech, writes that the borderline between primary and secondary connectives is not strictly defined. Some subcategories of secondary connectives act like primary connectives, and secondary connectives can even become primary through the process of grammaticalization (Rysová, 2015: 75). The conclusion is that primary and secondary connectives are complex categories with certain overlapping expressions. However, Rysová’s study concerns the Czech language; a similar study has not yet been carried out for English.

2.3.5.2 Modified, Parallel, and Conjoined Connectives

Apart from the distinction between one-word and multi-word connectives, the PDTB (Prasad et al., 2007) distinguishes between three types of more complex connective elements. Connectives can occur together with adverbs such as *only*, *even*, *at least*, etc. These multi-word connectives are called **modified connectives**, where the connective serves as the head of the

construction and is modified by an adverb (Prasad et al., 2007: 9). Examples of modified connectives that are annotated in the PDTB include the expressions *particularly since*, *even if*, *partly because*.

Parallel connectives are discontinuous pairs of connectives in which one token presupposes the presence of the other, and both together take the same two arguments. Parallel connectives include pairs such as *not only – but also*; *on the one hand – on the other hand*; *if – then*; *either – or*.

Conjoined connectives are expressions formed by two elements joined together by the coordinating conjunction *and*. Examples of conjoined connectives are *when and if* or *if and when*. In comparison to parallel connectives, conjoined connectives are continuous – their parts cannot be divided.

2.3.5.3 Contrastive, Collateral and Inferential Connectives

Fraser (1999), 946: divides discourse markers into two types: those that relate aspects of the explicit message conveyed by argument 2 with aspects of a message, direct or indirect, associated with argument 1; and those that relate the topic of argument 2 to that of argument 1.⁷ The first category includes contrastive markers, collateral markers, and inferential markers.

Contrastive markers signal that the explicit interpretation of argument 2 contrast with an interpretation of argument 1, and include connectives such as *but*, *however*, *although*, *in contrast*, *in comparison*, *conversely*, etc.

Collateral markers signal a quasi-parallel relation between argument 2 and argument 1 – the message of Arg2 parallels and possibly augments or refines the message of Arg1. Collateral markers can for example signal the addition of one more item to a list of conditions (*furthermore*), that the content of argument 2 is to be taken as the foremost exemplar of a concept (*above all*), or that there is a similarity between the arguments (*similarly*). Collateral markers include *and*, *also*, *besides*, *moreover*, etc.

Inferential markers signal that Arg2 is to be taken as a conclusion based on Arg1, or justification of Arg1. Examples of inferential markers are *so*, *of course*, *accordingly*, *as a consequence*, *then*, *all things considered*.

Fraser also distinguishes a fourth subclass of markers which serve to specify that Arg2 provides a reason for the content presented in Arg1. These include *after all*, *because*, *for this/that reason*, and *since*. All four categories of discourse markers above would be considered connectives by our definition.

⁷ The second type of discourse markers as defined by Fraser (1999) are not considered connectives.

The second category of discourse markers as described by Fraser (1999), called **topic change markers**, involves an aspect of discourse management:

(17) *This dinner looks delicious. Incidentally where do you shop?* (ibid: 949)

(18) *I am glad that is finished. To return to my point, I'd like to discuss your paper.*
(ibid: 949)

In the first example, *incidentally* signals that Arg2 is to be interpreted as a digression from the topic of Arg1, while in the second example, *to return to my point* signals the reintroduction of the previous topic of the discourse. It is the topic to which Arg1 is contributing, rather than its message, and these markers are not considered discourse connectives in our thesis. Topic change markers include expressions such as: *by the way, incidentally, just to update you, on a different note, that reminds me, to change to topic, to return to my point*, etc.

2.4 ENGLISH AS A FOREIGN LANGUAGE

2.4.1 Studies on Discourse Connectives in Texts of Non-Native Speakers

Discourse connectives have been the subject of many research papers in the last years. Researchers have focused on different groups of subjects, language registers and text types. Some studies are concerned with the use of connectives by native speakers, while others focus on the use of connectives by speakers of English as a foreign language (EFL), paying attention to instances of misuse and proposing possible methods of improvement in the instruction of non-native English speakers. This section makes reference to the studies by Aysu (2017), Chen (2006), Granger and Tyson (1996), Kwan (2017), Lahuerta Martinez (2004), Shaw (2009), and Tomešová (2017).

Aysu (2017) investigated the use of discourse connectives by Turkish speakers of English as a second language. Aysu analyzed a small-size corpus to determine the types and frequency of discourse connectives used. The results revealed a total of 180 discourse connectives, out of which 98 were instances of *and*, 51 of *but*, 18 of *because*. The connectives *then, so, also, too* and *still* appeared marginally. Using Fraser's (1999) taxonomy of discourse markers, 101 were found to be elaborative, 52 contrastive, 18 causative, and 9 inferential.

Chen (2006) examined the use of connectives in essays written by Taiwanese students, compared to essays by English speakers. The results of the study confirmed a higher density of connectives in the students' texts than in texts written by native speakers. The results also showed that the students used causal connectives (including *therefore* and *thus*) more frequently than native speakers. Chen (2006) pointed out that both groups, native and non-native speakers, relied on a small set of connectives in their writing. The occurrences of the ten most frequently

used connectives formed about two thirds of all the connectives. The most frequent connectives included *however*, *therefore*, *for instance*, *for example*, and *thus*. Chen (2006) discovered instances of misuses of the connective *therefore*, which some learners used to connect ideas that had no logical connection.

Tomešová (2017) worked on a comparison between the use of result/inferences connectives in corpora by native and non-native speakers of English. One corpus was formed by literary essays written by Czech university students of English and American Studies, and the other by journal articles of native expert writers. The thesis revealed a strong overall overuse of result/inference connectives by the students. It was also found that both types of writers generally preferred to place result/inference connectives initially. Further, students of English used connectives with immediate scope more frequently than native speakers.

Similarly, Shaw (2009) compared connectives (called *linking adverbials*) in texts written by native and non-native students. His results also showed a higher overall density of connectives in students' texts as compared to native English speakers. Shaw attributed this to the fact that students make more immediate connections between segments of text and use connectives to connect shorter and less developed propositions than native speakers. Shaw further found that the initial position of connectives was preferred, by both native and non-native speakers of English.

Lahuerta Martinez (2004) analyzed the use of discourse connectives in compositions by Spanish students. The main findings of her study were that students employed a limited variety of discourse markers, with some types used more frequently than others. The most frequently used type of discourse makers were elaborative markers (e.g. *indeed*, *in fact*, *actually*), followed by contrastive markers (e.g. *but*, *however*). She found a statistically significant relation between the ratings of the compositions and the number of discourse connectives used. The essays that contained more discourse markers obtained higher scores, and elaborative discourse markers were most closely related to the quality of writing.

Many researchers have focused on the identification of errors in the use of discourse connectives. Students tend to have difficulties with the use of connectives from the semantic, syntactic and stylistic aspects. The stylistic misuse of connectives has been discussed by researchers in detail (Chen, 2006; Granger and Tyson, 1996; Shaw, 2009). These researchers have been concerned with the incapability of some students to distinguish between the formal and informal uses of connectives such as *actually*, *anyway*, *besides*, *though*, etc. Connectives such as *actually* and *anyway* are typically confined to spoken language, and their use in formal

written language is inappropriate, yet speakers of English as a foreign language tend to use these types of expressions in written language as well.

Granger and Tyson (1996) also analyzed the use of stylistically inappropriate connectives and attributed the stylistic misuses of connectives by non-native speakers to the fact that students are not adequately taught stylistics in school, and that the differences between spoken and written language as well as formal and informal language are not paid attention to. In the case of native speakers, the misuses of connectives can be attributed to the transfer of expressions from conversation. Granger and Tyson's analysis revealed the wrong use of several connectives by English students, such as *moreover* and *on the contrary* (ibid: 22–23). Further, the excessive use of connectives is also considered erroneous and can have the consequence of incoherent and fragmented discourse (Chen, 2006: 125–126). Crewe (1990: 321) says that the overuse of connectives can be due the students' attempt to add surface logic to utterances in places where no deeper logical connections exist. He calls these cases of connectives “surface-level fillers”.

Already in 1942, Keating (1942) recognized that connective expressions (conjunctions, prepositions, adverbs) were the most difficult words to learn in a foreign vocabulary. He analyzed the difficulties of translation of connectives and showed that they caused many stops and hesitation. This might normally be attributed to the fact that transferring thought from one language to another is in itself a hard task and inevitably results in pauses. However, the observed group in his study did not lack knowledge of verbs and nouns, but rather of prepositions, adverbs and conjunctions. Keating (1942) argued that despite the importance of these linking items of discourse, they tend to be neglected by instruction. Connectives tend to be introduced only when they are specifically needed for composition purposes, and their meaning and use is only explained after. In the case of commonly occurring expressions such as *now* or *because*, this method is effective, since the speakers encounter them frequently. However, the same does not apply to other more complex structures. The students are expected to learn these expressions as they come, but they are not taught systematically. The difficulty of these words is also that they do not have a definite conceptual meaning and therefore do not exist in the brain independently. Keating (1942) wrote that abstract linking words should be paid attention to, or else students will have difficulty learning them properly. He suggested practicing tasks that would bring more attention to this class of words so that their meaning would be readily supplied to students later in the process of learning.

Kwan (2017) attributes the faulty use of connectives by Chinese speakers of English to the quality of English textbooks. Her paper critically evaluates the explanations, examples and

practices of the use of additive connectives in four different English grammar textbooks, concluding that the textbook explanations are simplistic and overlook the pragmatic and stylistic functions of individual additive connectives, and oversimplify their use in texts. The examples in the textbooks are artificial, and the use of connectives in them tends to be inappropriate or redundant, leading to students' misconceptions about the use and pragmatic implication of individual connectives.

2.4.2 The Common European Framework of Reference (CEFR)

Since this study concerns a specific level of language proficiency, let us focus on the description of the Common European Framework of Reference. The Common European Framework of Reference (CEFR, 2001) is a guideline used to describe the proficiency of second language learners. It was put together between the years 1989 and 1996 by the Council of Europe as part of the project *Language Learning for European Citizenship*. The CEFR was designed to facilitate comparisons between different systems of qualifications by describing the levels of proficiency required by existing standards, tests and examinations. It aims to provide a common basis for the elaboration of syllabi, examinations, textbooks, and other material important for second language learning and teaching across all European languages.

The CEFR describes language proficiency in terms of language competence, i.e. the necessary knowledge and skills that language learners must have to be able to use a language effectively for communication. The focus is not on what language learners have learned, but what they *can do* in the language. The CEFR provides descriptive scales of language proficiency for different communicative activities and strategies. **Communicative activities** include interactive activities such as conversation and correspondence, and the activities of speaking, writing, listening and reading. **Communicative strategies** are the means that a language user employs to perform the communicative activities mentioned above. Language proficiency is further described through **communicative language competences** (linguistic, pragmatic and sociolinguistic competences). Using descriptive scales in these categories, the CEFR divides speakers into six groups ranging from A1– C2, with A1 being the lowest language proficiency and C2 the highest. These groups, called **reference levels**, have become widely accepted as the standard across Europe. Below is a table of the levels and their denominations.

A		B		C	
Basic User		Independent User		Proficient User	
A1	A2	B1	B2	C1	C2
Breakthrough	Waystage	Threshold	Vantage	Effective Operational Proficiency	Mastery

Table 4: Table of Reference Levels in the CEFR (2001)

The reference levels offer a way to map the progress of learners as they improve their proficiency in a language. However, to be able to function as a basis for different languages, the framework is very general, and many categories are described vaguely. Even though each level is characterized in terms of the four abovementioned categories (speaking, listening, reading, writing), the analysis of concrete linguistic features in different linguistic situation almost lacking. Since the practical part of this study analyses the texts of B2 speakers, the following chapter involves a closer description of the B2 level in the CEFR.

2.4.3 B2 in the CEFR

On the general scale, the level B2 is described in the following way:

Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialization. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and Independent disadvantages of various options. (CEFR, 2001: 24)

Although this description attempts to define the reference level from different aspects of language use, it is very general and does not provide any information about the specific linguistic features that are expected to be used by B2 speakers. The CEFR contains two chapters that provide more detailed descriptions of the levels: Chapter 4 – *Language use and the language user*, and Chapter 5 – *The user's/learner's competences*. Chapter 4 contains descriptive **scales of communicative language activities and strategies**; chapter 5 contains **scales of general language competences** (e.g. the knowledge of the world, ability to learn, heuristic skills, etc.) **and communicative language competences** (linguistic, sociolinguistic, pragmatic competences). Chapter 4 divides communicative activities and strategies into

productive activities (speaking and writing) and receptive activities (listening and reading). Let us have a look at the description of written production in B2.

2.4.3.1 Written Production in B2

Here, illustrative scales are provided for three categories: overall written production, creative writing, and reports and essays. The table below shows the description of B2 speakers in each of these categories:

Overall written production	Creative writing	Reports and essays
Can write clear, detailed texts on a variety of subjects related to his/her field of interest, synthesizing and evaluating information and arguments from a number of sources. (CEFR, 2001: 61)	<p>B2+:⁸ Can write clear, detailed descriptions of real or imaginary events and experiences, marking the relation between ideas in clear connected text, and following established conventions of the genre concerned.</p> <p>B2: Can write clear, detailed descriptions on a variety of subjects related to his/her field of interest. Can write a review of a film, book or play. (CEFR, 2001: 62)</p>	<p>B2+: Can write an essay or report which develops an argument systematically with appropriate highlighting of significant points and relevant supporting detail. Can evaluate different ideas or solutions to a problem.</p> <p>B2: Can write an essay or report which develops an argument, giving reasons in support of or against a particular point of view and explaining the advantages and disadvantages of various options. Can synthesize information and arguments for a number of sources. (CEFR, 2001: 62)</p>

Table 5: B2 Written Production in the CEFR (2001)

The description of writing in B2 focuses more on the content rather than the lexical or grammatical means of expression. Only the creative writing scale contains some information on the presence of linking elements in text: “... marking the relation between ideas in clear connected text, and following established conventions of the genre concerned,” but it does not mention any specific linguistic devices that are used to achieve this.

⁸ In the CEFR descriptors, a distinction is made between the ‘criterion levels’ (e.g. B2 or B2.1) and the ‘plus levels’ (B2+ or B2.2). The plus levels represent higher language proficiency. In B2, the plus level (B2+) is called ‘strong vantage’ (CEFR, 2001: 32).

2.4.3.2 Discourse Competence in B2

Chapter 5 contains three sections of learners' communicative language competences: linguistic competences, sociolinguistic competences and pragmatic competences. The description of each subcategory of linguistic competences contains a list of features used for evaluation. However, these features are not described or specified in relation to the individual reference levels. In the scale of grammatical accuracy, B2+ speakers are said to have “minor flaws in sentence structure” (CEFR, 2001: 114), but it is not stated what kind of mistakes they make. In terms of orthographic control, B2+ speakers “can produce clearly intelligible continuous writing which follows standard layout and paragraphing conventions” (CEFR, 2001: 118), but it is not explained what constitutes such “clearly intelligible continuous writing”.

The most relevant information in connection to discourse connectives is provided in Chapter 5 in the section about discourse competence. **Discourse competence** is described as “the ability of a user/learner to arrange sentences in sequence as to produce coherent stretches of language” (CEFR, 2001: 123). According to the CEFR, this includes the ability to structure and manage discourse in terms of thematic organization, coherence and cohesion, logical ordering, style and register, and rhetorical effectiveness. Below is a descriptive scale of coherence and cohesion across all reference levels:

	COHERENCE AND COHESION
C2	Can create coherent and cohesive text making full and appropriate use of a variety of organizational patterns and a wide range of cohesive devices.
C1	Can produce clear, smoothly flowing, well-structured speech, showing controlled use of organizational patterns, connectors and cohesive devices.
B2	Can use a variety of linking words efficiently to mark clearly the relations between ideas. Can use a limited number of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some ‘jumpiness’ in a long contribution.
B1	Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.
A2	Can use the most frequently occurring connectors to link simple sentences in order to tell a story or describe something as a simple list of points. Can link groups of words with simple connectors like <i>and</i> , <i>but</i> and <i>because</i> .
A1	Can link words or groups of words with very basic linear connectors like <i>and</i> or <i>then</i> .

Table 6: Descriptive Scale of Cohesion and Coherence in the CEFR (2001)

Although the CEFR provides a separate scale on coherence and cohesion, the only mention of specific discourse connectives appears in the description of the two basic levels, A1 and A2. B2 speakers are said to use a “limited number of cohesive devices”, but it is not mentioned which devices these are in comparison to the other levels. B2+ speakers are said to use a “variety of linking words”, but again, a detailed linguistic analysis is not present. In contrast to the B2 level, speakers with a higher language proficiency are said to show a “controlled use of connectors and cohesive devices” (C1), and speakers at the highest level of proficiency should be able to implement a “wide range of cohesive devices” (C2). It becomes clear that, the higher the level of proficiency in a language, the better the use of cohesive devices. The CEFR even mentions specifically that when learning a foreign language, a speaker is likely to start with short segments, usually single sentences, and then proceeds to more complicated structures.

However, even though the CEFR stresses the importance of discourse competence, especially in higher levels of proficiency, it does not explain the differences in the frequency and use of cohesive devices in the individual reference levels. Discourse connectives are key in indicating semantic relations in text and “arranging sentences in sequence” as the CEFR itself states, but they are mentioned only cursorily. This thesis aims to take a closer look at discourse connectives used by B2 speakers and to provide a more detailed description of this level within the European Framework.

3 MATERIAL AND METHOD

3.1 Subjects

This study featured 161 subjects, all speakers of English as a second language (ESL). At the time of the collection of data, all subjects were actively studying at the University of Economics in Prague (VŠE, Vysoká škola ekonomická). The students were both male and female participants, aged between 20–30 years. Since the research was anonymous, we do not have much detailed information about the students. Some students included more specific information about themselves in the texts they wrote, but the type and extent of information they included differed from student to student. Factors such as their background, specific area of study, prior education in English, contact with native English speakers, knowledge of languages other than English, experiences from studying abroad, etc., could all play a role in their level of English. Therefore, a normalized test was required to see whether the group was homogeneous in terms of language proficiency.⁹

Prior to the collection of data, all students were given an entry test to prove that their English knowledge corresponded to the level B2 according to the Common European Framework. The results of the tests and the language proficiency were evaluated by a committee of teachers of English as a foreign language, who were all trained in the assessment of language proficiency according to the criteria given by the CEFR, and who have long-term experiences with evaluating students' work. The committee followed the descriptive scales of the CEFR in areas of grammar, vocabulary and writing. After passing the test, the students were given the assignment to write a short formal letter as an application for the Erasmus program. Although they were all classified as B2 speakers, there were slight differences in the language proficiency of individual speakers, which is an aspect that cannot be avoided. Some speakers were able to use more complex structures and vocabulary than others. The difference was for example in the use of longer sentences, complex/compound sentences, variety of the types of subordinate clauses, richness of vocabulary, frequency of grammatical mistakes, and, ultimately, the use of connectives. However, these differences were minor in relation to the reference level and were considered as differences *within* the level B2, or between the categories of *vantage* (B2/B2.1) and *strong vantage* (B2+/B2.2).

⁹ We faced an alternative to compare the B2 texts with those of more advanced speakers of English (level C1). However, due to the extent of this analysis and factors like the area of study (all C1 speakers were university students of English, educated in linguistics), we decided to exclude the C1 texts from the analysis.

3.2 Material

Since this thesis concerns the written language of non-native speakers, the material used for the analysis are short written texts. The total number of texts corresponds to the total number of speakers, 161. Due to the length of the analyzed material, all original texts in their full extent are available only as the electronic appendix of this thesis in the online depository of the Charles University. The list of examples that were chosen for closer analysis is in the Appendix.

The subjects were asked to write a short formal email on a very specific topic – an application letter for the Erasmus student exchange program. The texts were supposed to have a minimum length of 130 words and maximum of 150 (which constitutes approximately 12 lines of text). In this short text, the students were supposed to express their motivation and reasons to study abroad. The exact assignment that each student was asked to complete was the following:

Formal Email Writing: Application for the Erasmus student exchange program

You would like to apply for the Erasmus student exchange program at a European university of your choice. You are planning to spend a semester there. Write your motivation email. Use formal English, write 130-150 words.

Remember to follow all the instructions below and use proper email structure, sentence structure, vocabulary, grammar, and spelling. Please revise and proofread your email before you hand it in.

Please make sure to include the following points:

- *Your name (real or a pseudonym)*
- *The reason for writing*
- *Where you are studying now and what exactly you are studying*
- *Why you are interested in the particular institution*
- *A summary of relevant work/study experience, courses, training, projects, and language skills*
- *How you will benefit from your participation in the program*

Table 7: B2 Assignment – Erasmus Application Letter

As can be seen, the structure of the email was very strictly given. The students were first asked to introduce themselves, then they were supposed to list their reasons for writing, their field of study, the motivation to take part in the exchange program, a summary of their achievements so far, and the ways in which the program would be beneficial to their future career. The text type showed to play a role in the choice of connectives, especially in the types of semantic relations expressed (most frequently expansion and cause).

What also played a role was the instruction to write in the formal register, which is bound to certain kinds of expressions (as opposed to informal or spoken language). At the same time, the fact that the task was assigned to the students and the application letters were not meant for actual use might have affected the speakers' choice of vocabulary and diminished their motivation in writing. For example, even though the students were asked to revise their text before handing it in, many of the emails contained grammatical mistakes. These mistakes concerned mainly spelling, the use of tenses and verb forms, inadequate use of vocabulary, and punctuation. Most of these errors were due to the speakers' intermediate level of English, but some of them were inadvertent mistakes that could be avoided by simple proofreading. Nevertheless, all collected samples were considered in the analysis. The electronic appendix contains all texts in their original form.

There were slight differences in the length of the letters. Most texts were 5–12 sentences long, but some texts were slightly longer, containing more than 12 sentences (texts number 9, 40, 72, 73, 87, 91, 109, 130, 143, and 159). A few texts, on the other hand, were considerably shorter, containing less than five sentences (texts number 59, 116, 147, and 149). These differences in length were considered – the analysis contains absolute quantities, but also average values with relative proportions of data.

3.3 Data Processing

Due to the complexity of the topic of discourse connectives, rather than running automatic searches and scripts, most of the analysis was done manually, from sentence to sentence. All texts were first copied into one document to form a corpus and then numbered. Then, the connectives in each text were determined and annotated individually. The structural uses of *and* were also annotated right away. For a general quantitative analysis, the number of connectives in each text and the number of sentences in each text was calculated. We also calculated the average number of the structural use of *and* in each text.

During a second revision, we analyzed several different categories for discourse connectives and discourse relations, namely types of connectives (one-word/multi-word, primary/secondary connective), realization form (word class or syntactic structure), position of connective in the sentence (initial/post-initial/pre-medial/medial/final), inter-/intra-sentential scope, type of arguments (verbal/nominal), semantic types of relations according to the PDTB (Prasad et al., 2007). The main program we worked with to process the data was Microsoft Excel. For a demonstration, one sample of a B2 text used for the analysis and the respective segment from the table of annotations in Microsoft Excel are shown below.

13

Hello, recently I've decided to participate in the Erasmus program. I think that's a very big opportunity, **not just** to improve my language skills, **but also** to try to live abroad for a particular time separated from a family, meet many new people, discover new place **and** what's for me the greatest point, try to live a different style of living, **because** live is a change **and** I, living whole my life at Prague, the life here reminds me kind of a stereotype. **So** this opportunity really went forward to me.

I study the University of Economics in Prague, Business faculty. **If** I was chosen for the program I would **also** be interested in a part time job there, **because** I would like to make any money beside the studying **and also** could involve that experience into my CV.

Looking forward to hearing from you,

XY

Table 8: Example of B2 Text from the Student Corpus (Text No. 13)

Text no.	Connective	Type of Arg.	Semantic Type of Relation	Position, intra-/inter-sentential	One-word/multi-word, primary/secondary	Morph./syntactic Class
13	not just...but also	infinitive	expansion, conjunction	post-initial+medial ¹⁰ , intra-	multi-, primary	correlative conjunction (conj+adverb)
13	and	infinitive	expansion, conjunction	medial, intra-	one-, primary	coordinating conjunction
13	because	verbal	cause, reason	initial, intra-	one-, primary	subordinating conjunction
13	and	verbal	expansion, conjunction	medial, intra-	one-, primary	coordinating conjunction
13	So	verbal	cause, result	initial, inter-	one-, primary	adverb
13	If	verbal	condition, hypothetical	initial, intra-	one-, primary	subordinating conjunction
13	also	verbal	expansion, conjunction	post-initial	one-, primary	adverb
13	because	verbal	cause, reason	initial, intra-	one-, primary	subordinating conjunction
13	and also	verbal	expansion, conjunction	medial, intra-	multi-, primary	coordinating conj+adverb

Table 9: Example of Connective Annotation in Microsoft Excel (Text No. 13)

¹⁰ In correlative pairs, the first part of the pair is found in argument 1, and the second pair in argument 2. Here, the position of the first part was post-initial, of the second medial.

Each discourse connective was annotated separately for all the categories listed in the columns above. Some categories (e.g. one-word/multi-word, position) were quick to determine, while others (e.g. semantic type of relation, scope) had to be analyzed in relation to the surrounding context. In total, there were 1011 connectives in the 161 texts by B2 speakers. Since this thesis concerns the analysis of the language of non-native speakers, every expression that the speakers used to link segments of text and signal a semantic relation was marked as a connective. As will be explained in detail in the practical analysis, not all connectives were used correctly, and some were used to express different kinds of relations at once.

Using the functions of Microsoft Excel, we calculated the absolute and average values in different categories. For a better visual representation, we used Microsoft Excel to create graphs of the results. We used either bar charts, line charts, pie charts or point graphs, depending on how we wanted the data to be displayed. For example, bar charts were most appropriate when the analysis included many subcategories, while pie charts were useful to show the relative percentages of values.

The method chosen for the practical part of this study was not unproblematic. The fact that most of the analysis was done manually resulted in complications when data needed to be changed, which affected the data in all other categories. Also, work done by a person can never be flawless. However, all data was checked to eliminate typographical errors, and the classification of connectives was thoroughly revised.

3.4 Hypothesis

Basing our hypothesis on previous studies of the topic and the description of the level B2 in the Common European Framework of Reference, we have reason to assume that B2 speakers will use a limited number of connectives, most of which will be simple one-word primary expressions realized by the word classes of conjunctions, adverbs, or by prepositional phrases. Due to their grammatical features, discourse connectives will tend to appear medially and will link verbal arguments¹¹. Their scope will be restricted, they will link arguments within the sentence. The main type of relation expressed by connectives will be additive, since it is the most straightforward way to link arguments to preceding context. We also assume that the frequency of connectives per letter will not be too high, since the speakers will predominantly use other means of cohesion to join utterances.

¹¹ For instance, coordinating conjunctions are medial (and we assume a prevalence of paratactic relations); conjunctions and adverbs do not allow for the nominalization of arguments they introduce.

4 ANALYSIS

4.1 General Findings

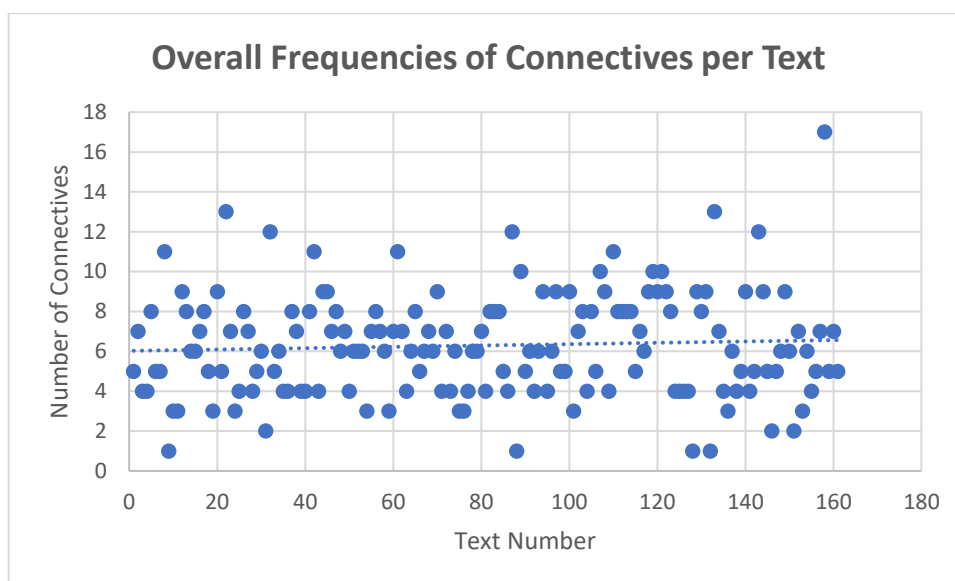


Figure 2: Overall Frequencies of Connectives per Text

The practical part of this study includes both a quantitative and qualitative analysis of the data. Let us first begin with an overall analysis of the frequency of connectives, disregarding the specific types of expressions (whether one-word/multi-word, initial/post-initial/medial/final, primary/secondary, etc.). The diagram above shows an overview of the frequencies of connectives per text. The vertical axis marks the total number of connectives per text and the horizontal axis marks the text number. This means that each blue dot on the chart stands for one text, and depending on the position on the vertical axis, the number of connectives in that text. As can be seen, each text contained at least one connective. The minimal number of connectives per text was 1 (which appeared in 4 texts), and the maximum was 17 (in only 1 text). This produces a relatively large range of 16 connectives per text ($17 - 1 = 16$, meaning that the values of connectives/text oscillated between 1 and 17). As can be seen on the graph, the data is dispersed around the number 6. This trend is represented by the dotted horizontal trendline. The trendline is a moving average line that joins the individual calculated averages of connectives per text. To calculate the overall average of connectives per text, we divided the total number of connectives (1011) by the total number of texts (161). This gave us a mean value of 6,28 connectives per text.

Although this data gives us an idea about the frequency of connectives per one text, it disregards the number of sentences per text. Although all texts were 130–150 words long, they differed in the number of sentences. Most texts contained 5–12 sentences, but some texts were

slightly longer, containing more than 12 sentences (texts number 9, 40, 72, 73, 87, 91, 109, 130, 143, and 159). A few texts, on the other hand, were considerably shorter, containing less than five sentences (texts number 59, 116, 147, and 149). Since the sentence is regarded as a basic unit of discourse (see 2.1.5), we decided to count the number of sentences per text to achieve more precise data on the frequency of connectives. Using the data about the frequency of connectives per text and the frequency of sentences per text, we calculated the ratio of connectives per sentence for each text. For instance, in a text with 5 connectives and 8 sentences, the ratio was $5/8=0,63$. Applying these values, we then calculated the average number of connectives per 1 sentence, getting a value of 0,727 (connectives/sentence). For the opposite ratio, or the average number of sentences per one connective, we calculated a value of 1,796. This means that on average, there were 1,8 sentences per 1 connective. A table of the first five annotations used for calculating the mean values is shown below.

A (connectives per text)	B (sentences per text)	C (Ratio A/B)	D (Ratio B/A)
5	8	0,63	1,60
7	9	0,78	1,29
4	7	0,57	1,75
4	8	0,50	2,00
8	11	0,73	1,38

Table 10: Average Ratios of Connectives per Sentence and Sentences per Connective

The average ratios were calculated using the “scalar product” function in Microsoft Excel. The formula used for column C (ratio A/B) was $(A2:A162;1/(B2:B162))/161$, and the formula used for column D (ratio B/A) was $(B2:B162;1/(A2:A162))/161$.

The most frequent number of connectives per application text, or the mode value of the data, was 4. This is illustrated in the following diagram, where the vertical axis stands for the number of connectives per application text, and the horizontal axis stands for the frequency of these numbers. For example, starting from the base of the vertical line, the frequency of 1 connective per text appeared in 4 texts, the frequency of 2 connectives per text appeared in 3 texts, of 3 connectives in 11 texts, etc. As stated before, the most frequent number of connectives per text (the mode value) was 4, with a total of 28 occurrences. Some frequencies (e.g. 13, 14, 15, 16) were not found at all. The value of 0 also does not appear, which proves that each text contained at least 1 connective.

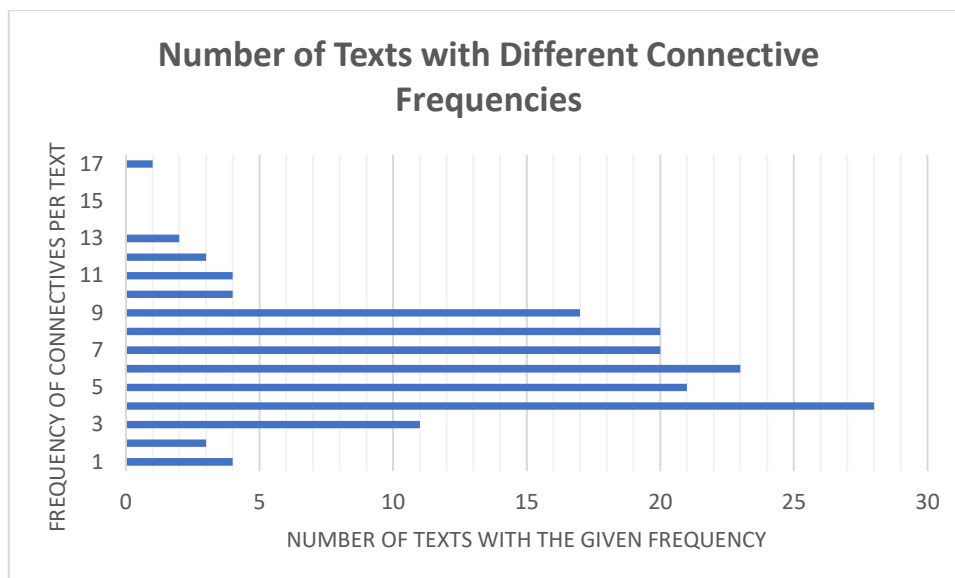


Figure 3: Number of Texts with Different Connective Frequencies

For further qualitative analysis, we looked at the different categories of discourse connectives and discourse relations. The analyzed categories were: morphological and syntactic classes of connectives, one-word and multi-word connectives, primary and secondary connectives, initial/post-initial/medial/final connectives, intra-/inter-sentential connectives, types of discourse arguments (verbal/nominal), and semantic types of discourse relations.

4.2 Morphological and Syntactic Classes of Connectives

We worked with the hypothesis that most connectives would be simple one-word expressions that belonged to the morphological classes of conjunctions, adverbs and prepositional phrases. The analysis showed that the most frequent morphological class were indeed conjunctions (*and, but, because*, etc.), with a total number of 675. The second most frequent class were adverbs (*however, therefore*, etc.), with a total of 208. Interestingly, the third most frequent class were clauses (e.g. *That is the reason why*, see 4.2.4), with a total of 61 occurrences. 55 connectives were realized by prepositional phrases. Below is a graph that shows these values in relative percentages. Conjunctions formed 69% of the occurrences, adverbs 21%, clauses 6% and prepositional phrases 4%.

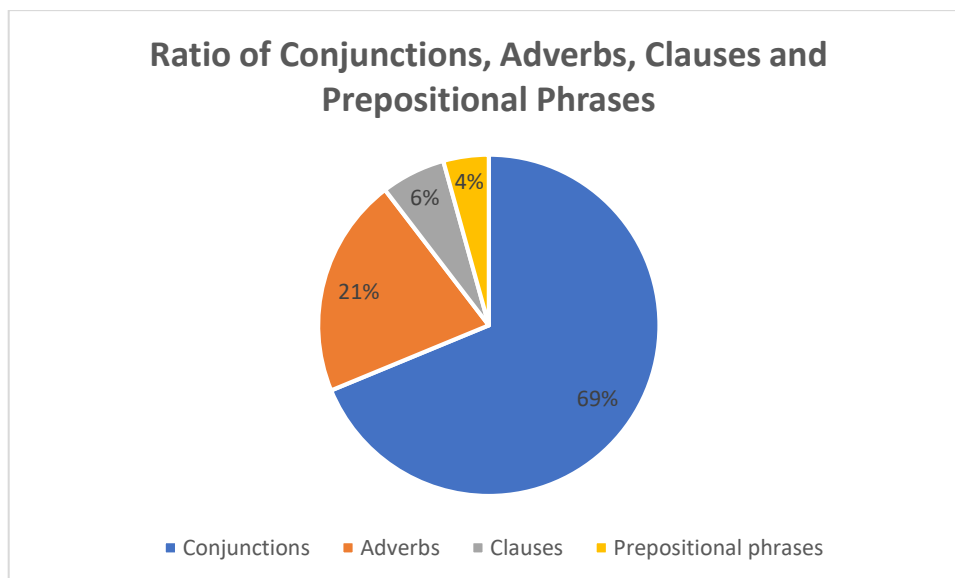


Figure 4: Morphological and Syntactic Classes of Connectives

The following chapters include a more detailed analysis of the individual morphological and syntactic classes, looking not only at quantitative values, but also at individual examples from the texts.

4.2.1 Conjunctions

Out of the 675 conjunctions, 454 were coordinating and 221 were subordinating, which constitutes a ratio of roughly 2:1, as depicted by the graph below.

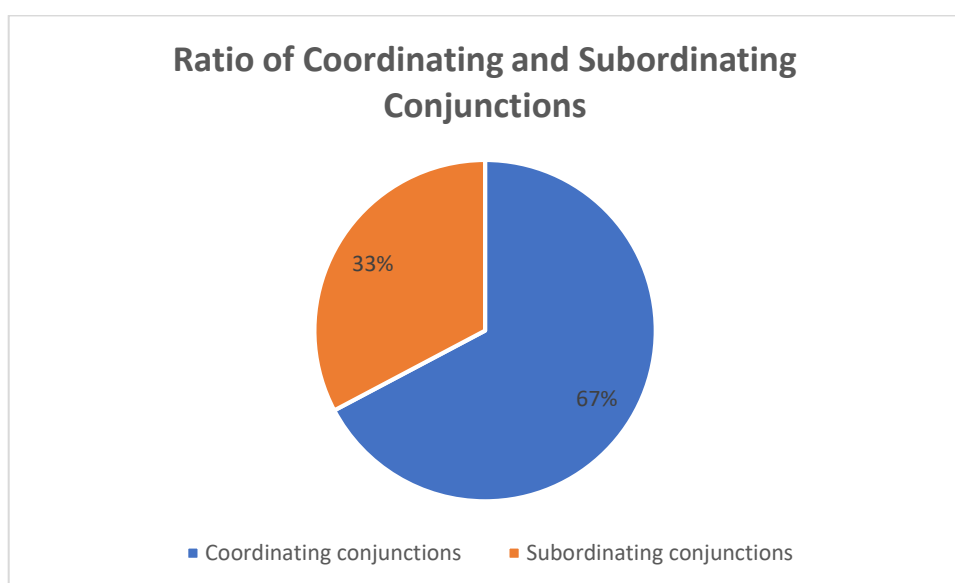


Figure 5: Ratio of Coordinating and Subordinating Conjunctions

The analysis of the occurrences of individual connectives showed certain trends. There was an overall strong prevalence of the coordinating conjunction *and*. This conjunction was, in fact, the most frequent discourse connective of all (including other morphological and syntactic

classes), with a total of 372 occurrences. Below is a graph demonstrating the total frequencies of the four coordinating conjunctions, *and*, *but*, *or* and *for*.

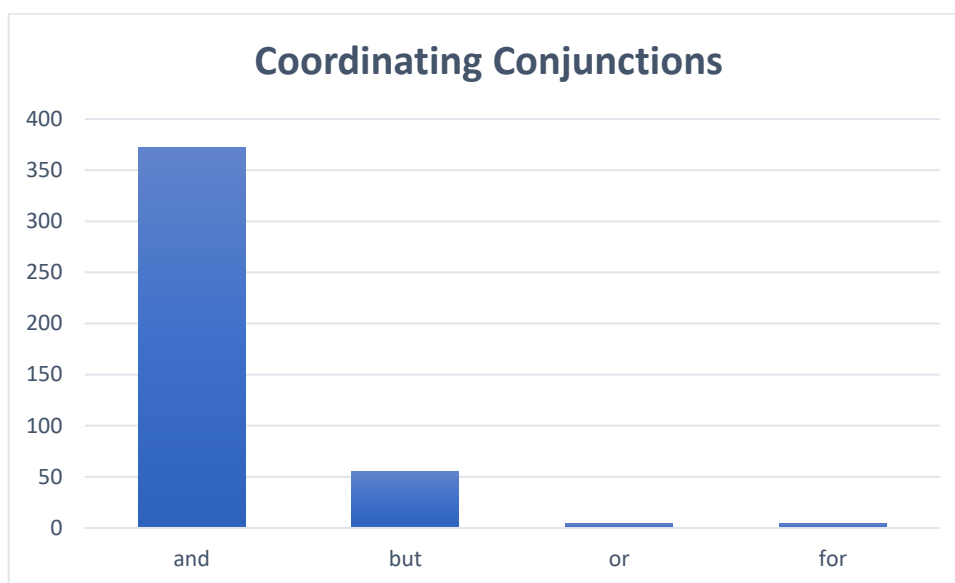


Figure 6: Coordinating Conjunctions in B2 Texts

As can be seen, there was a marked difference between the use of *and* and the other three coordinating conjunctions. The contrastive conjunction *but* appeared 55 times, the alternative *or* and causal *for* each appeared only 4 times. The range of subordinating conjunctions used was relatively limited. The subordinating conjunctions used were: *because*, *if*, *since*, *after*, *although*, *when*, *while*, and *before*. A diagram with the frequencies of these individual conjunctions is shown below.

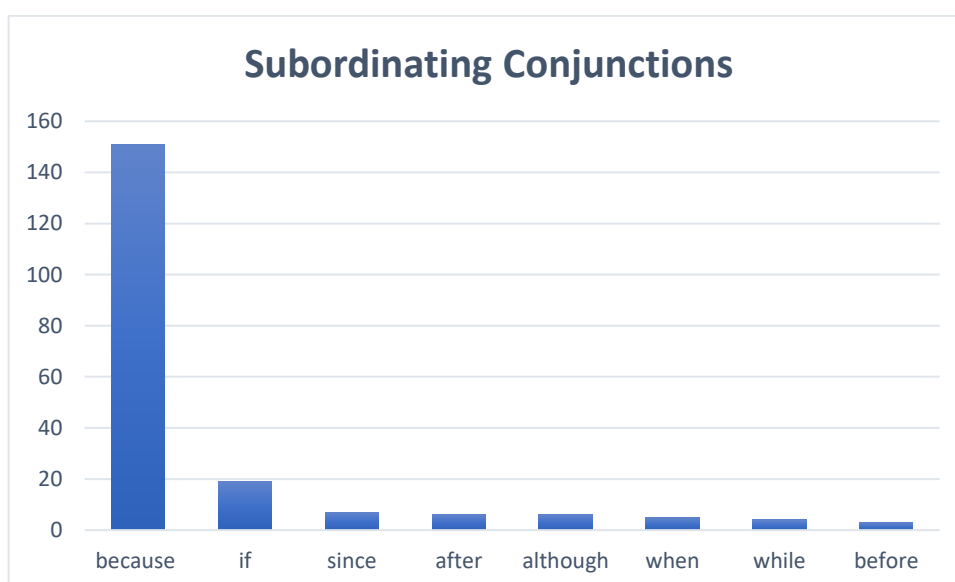


Figure 7: Subordinating Conjunctions in B2 Texts

The most frequent subordinating conjunction by far was *because*, with a total of 151 occurrences. The conjunction *if* had 19 occurrences, *since* had 7. *After* and *although* appeared 6 times, and *when*, *while* and *before* less than five times in the texts.

4.2.1.1 Cohesive vs. Structural “And”

Since *and* was the most frequent connective of all, we focused on a more thorough analysis of the conjunction. Halliday and Hasan (1976) differentiate between the structural and connective use of the connective (see 2.3.2.12.1.5.2). In the texts by B2 speakers, the coordinating conjunction appeared 523 times in total, but in 151 cases, its role was only structural. To explain the structural uses of *and*, consider the following examples from the B2 texts:

(1) *I can speak German, English **and** Spanish.*

(2) *I am studying Business **and** Management at University of Economics in Czech Republic.*

In both examples above, the role of *and* is the coordination of noun phrases at a sentential level. The order of the elements (noun phrases) can be easily swapped without changing the meaning of the utterance: *I can speak German, Spanish and English / I can speak English, German and Spanish / I can speak Spanish, English and German*, etc. Further, the items joined by coordination function as a single complex element of structure. In both examples, this element is the direct object of the sentence. If more elements were added to the list, they would still constitute a single clause element:

I (S) can speak (V_{modal}) [German, English, Spanish, Czech, Italian **and** French] (O_D).

Such examples were therefore not annotated as connectives. A line chart of the structural occurrences of *and* in the texts is shown below.

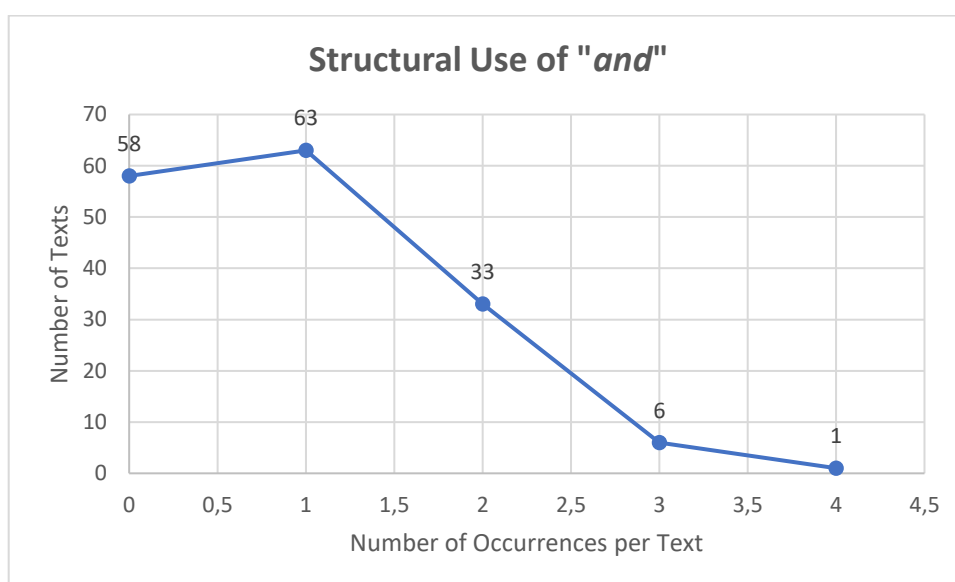


Figure 8: Structural Use of “and” in B2 Texts

In total, 58 texts contained no structural uses of *and*, 63 contained 1, 33 contained 2, 6 contained 3, and 1 contained 4. This means that most connectives contained at least 1 structural use of the

coordinating conjunction. However, much more frequently, *and* was used in its cohesive role, such as in the following example:

(3) *Dear Mr. Novak, my name is XY **and** I am writing to you because I heard about your Erasmus offer in Brighton, England.*

In this example, the clauses cannot be rearranged without losing cohesion and logical thematic development; they follow one another as the text unfolds and the speaker reveals more information about himself/herself and his/her reasons for writing. The coordinating conjunction *and* expresses a dependency between the two propositions that are realized by finite clauses, or verbal arguments (*Arg1 = my name is XY; Arg2 = I am writing to you...*). Its meaning extends over the entire sentence. The semantic relation between the two propositions is additive – the second argument adds new information to the first. The second proposition then introduces a causal dependency (*I am writing to you **because** ...*). This structure, “My name is XY and I am writing to you because ...”, was used quite frequently in the texts. It appeared in 18 out of the 161 texts, which forms a percentage of about 11,2% ($18/161 \times 100$). This percentage means that 1 in every 11 texts began with the chunk. The frequency of occurrence of the chunk is probably due to the instructions that the speakers were given, since they were explicitly told to introduce themselves and write their reason for writing.

In many texts, the coordinating conjunction *and* was used excessively. For example, in text number 87, out of 9 connectives, 6 were cases of *and*, which constitutes two thirds of the total count. The entire text with all occurrences of the additive *and* is shown below.

87
Dear all,
My name is XY and I am writing you because I would like to participate in Erasmus at Cambridge University.
I am studying in Prague, where I also live and I am studying first year of University of Economics.
I am interested in Cambridge University because a lot of my friend recommended it to me. I had been working from my age of fifteen. I used to do bartender, IT technician and now I program applications in Java. This summer I made my trip to Brighton, I arrived and have set up just one week of staying in some lady's house, no job. In the end I was renting a flat in center of Brighton and I was working as bartended in one of the best bars on the Brighton coast. I made quite a lot of money and improved my English a lot. When I was younger I also participated some language camps in Czech Republic, but it is not even worth of mentioning.
I will benefit most in improving my English, learning some experience, experiencing English education system and meeting new people.

Figure 9: Example of B2 Text (No. 87) with Cohesive Use of "and"

The text above contains 10 sentences and 9 connectives, which gives a ratio of roughly 1 connective per sentence. In several of these sentences, the use of *and* is unnecessary and could be omitted to achieve a more formal register. For example, the sentence

- (4) *I am studying in Prague, where I also live and I am studying first year of University of Economics,*

which contains three finite clauses, could be condensed into a much simpler and clearer utterance, without the use of *and*:

I am studying the first year of University of Economics in Prague, where I also live.

In some cases, the use of *and* was rather superfluous, and could be replaced by a full stop or comma:

- (5) *I would like to thank you for giving such an opportunity for young generation and I hope my dream will come true.*

- (6) *Thank you for your time and I will be looking forward to hearing from you.*

- (7) *I would like to travel on my own and lastly I would like to be a part of your institution.*

Instead of the coordinative *and*, some speakers used asyndetic coordination (coordination by means of punctuation) to express additive relations. Most of the clauses in the following compound sentence are joined by commas:

- (8) *Firstly, I would like to introduce myself, my name is XY, I am 23 years old and I am studying at the Economics University in Prague, our main study program is based on the Business Administration.*

Similarly, the author of text 71 used full stops between arguments, resulting in a sequence of short coordinated sentences with no overt marking of semantic relations:

- (9) *My name is XY. I am 22 years old. At the moment I am studying Accounting at the University of Economics in Prague.*

Although the relations between the segments are not overtly expressed, the sentences are semantically connected and constitute a text. The relations of expansion are implicit (*Arg1 – and – Arg2 – and – Arg3*) and can be deduced from the structure and semantics of the utterance. The cohesive means that are applied to create texture are not discourse connectives, but relations of reference (in specific, personal reference using the personal pronouns *my*, *I*).

Generally, there was an overuse of the coordinating conjunction *and*. It was used not only to express additive relations, but also to chain segments of discourse that were not directly related to each other. *And* was mostly used as a connective, though in 151 cases its use was only structural.

4.2.1.2 Subordinating “Because”

As mentioned in the overview (see 4.2.1), the most frequent subordinating conjunction used by the B2 speakers was *because*. The high frequency of occurrence might be attributed to the type of text used for this analysis – the application letter. The purpose of an application letter is to introduce oneself and explain one’s motivation and reasons for applying. Since *because* is a primary conjunction that expresses exclusively relations of cause (reason), its use was very recurrent. The following sentence is a prototypical example of the structure *Arg1 (statement) – because Arg2 (reason)*, used very often by the speakers:

- (10) *I would like to apply for this exchange program **because** it is interesting for me to get this experience.*

The speakers did not seem to have trouble using the conjunction *because* in the middle of the sentence, but some grammatical mistakes were found when *because* was placed at the beginning of the sentence. The instances where *because* was used at the beginning of the sentence are shown in the following table. The argument introduced by *Because* is on the right, and the preceding linguistic context that contains the first argument is on the left.

Text no.	Argument 1	Connective	Argument 2
7	I think that relevant work is not important like happy from new experience in charity.	Because	“free help” is goodness.
22	I really enjoy a participation in these projects.	Because	I think every culture has something special
22	So I think that one of the biggest benefit for me (...) would be finding new friends (...) and an opportunity for good job.	Because	I really like travelling
119	I decided to apply for the Erasmus Student Exchange program.	Because	I’d like to improve my language skills
133	I can study well	Because	I have nice results from high school

Table 11: Initial Use of “Because” in B2 Texts

In all the examples above, *Because* is used to link the sentence it introduces with the preceding context. It expresses a causal relation, where *Arg1 – Because Arg2*. However, its use at the beginning of the sentence is grammatically incorrect and leads to a loss of coherence – there should be no full stop between the two arguments. The only occurrence where the initial placement of *Because* was correct was in the following example:

- (11) ***Because** I am student of economics and management, I found suitable host university in Paris for me.*

Here, *because* does not link the sentence to preceding context, but expresses a reason for the argument that follows. The structure of the utterance is ***Because*** *Arg1 (reason)* – *Arg2 (result)*.

To sum up, *because* was the most frequent subordinating conjunction, with a total of 151 occurrences. *Because* was used to express exclusively relations of cause (reason). It was used mostly in the middle of the sentence, but sometimes incorrectly at the beginning of the sentence, which led to a loss of coherence.

4.2.1.3 Correlative Pairs

8 of the subordinating conjunctions were correlative pairs, which are labeled by the PDTB (Prasad et al., 2007) as *parallel connectives*. The typical feature of parallel connectives is that the first expression in the pair is part of argument 1 and the second expression is part of argument 2. The following table shows a list of all correlative pairs found in the texts.

Text no.	Connective	Semantic Type of Relation	One-/multi-word, primary/secondary
13	not just...but also	expansion, conjunction	multi, primary
16	not only...but also	expansion, conjunction	multi, primary
17	not just...but also	expansion, conjunction	multi, primary
18	- ...but also	expansion, conjunction	multi, primary
42	not only...but as well	expansion, conjunction	multi, primary
63	not only...but also	expansion, conjunction	multi, primary
131	either...or	expansion, alternative	multi, primary
146	either...or	expansion, alternative	multi, primary

Figure 10: List of Correlative Pairs (Parallel Connectives)

As can be seen in the list above, most pairs express a relation of conjunction (*not only X, but also Y*). This relation is expressed using different lexical items (*not just/not only X, but also/but as well*), but the semantics remains the same. In text number 18, the first expression of the correlative pair is missing, but this is probably an inadvertent mistake on the side of the speaker. The sentence that contains the connective is as follows:

*This will be **not only** a great investment in my education, **but also** a new source of new great experiences and memories in the capital of Denmark.*

Here, the first part in the pair (*not only*) and the relation of expansion between the two arguments are implied. The second type of relation expressed by correlative pairs was alternative and appeared in two examples (texts 131 and 146), as shown in the table above.

4.2.1.4 Other Multiple Conjunctions

As described in chapter 2.3.5.2, the PDTB (Prasad et al., 2007: 9) distinguishes three special cases of multi-word connectives: *parallel*, *modified* and *conjoined*. Modified connectives are defined as connectives that occur with adverbs such as *only*, *even*, *at least* and so on. The examples of modified connectives that the PDTB lists are *particularly since*, *even if* or *partly because*, where the connective (*since*, *if*, *because*) is the head of the construction and the adverb (*particularly*, *even*, *partly*) is the modifier. Such a construction appears in text number 41. The connective is in **bold** and the argument it introduces is underlined:

- (12) *The reason why I want to attend this Erasmus project is **mainly because** I would like to improve my english or spanish or learn a new language if it is possible.*

Here, the head of the construction is the subordinating conjunction *because*, which is pre-modified by the adverb *mainly*. Another modified connective appears in text number 55:

- (13) *I am very interested in this project **especially if** I will attend your university.*

Here, the subordinating conjunction *if* is modified by the focusing adjunct *especially*. In terms of meaning, however, the connective is used improperly.

There were 8 other cases of multiple constructions which were not as easily defined. These were the constructions *and also*, which appeared 7 times, and one case of the construction *and exactly*. A table of these constructions is shown below.

Text no.	Connective	Type of Argument	Semantic Type of Relation	Position
74	and also	verbal	expansion, conjunction	medial
77	and also	verbal	expansion, conjunction	medial
80	and also	verbal	expansion, conjunction	medial
89	and also	verbal	expansion, conjunction	medial
89	and also	verbal	expansion, conjunction	medial
96	and also	verbal	expansion, conjunction	medial
96	and also	verbal	expansion, conjunction	medial
97	and exactly	verbal	expansion, conjunction restatement, specification	medial initial

Table 12: Examples of Multi-Word Connectives Used by B2 Speakers

The constructions resemble modified connectives, since they also combine the use of a coordinating conjunction (*and*) and an adverb (*also/exactly*). However, the order of the elements in the construction is inverted. To determine whether these were truly examples of modified connectives, we looked at the semantic properties of the words in the expressions. In the case of *and also*, both words in the multi-word construction indicated a relation of expansion

(conjunction). We tried to determine the head of the construction by a substitution test. First the construction *and also* was replaced by the coordinating conjunction *and*, and then by the adverb *also*.

- (14) *It could give me worth experience **and also** help me in my future career*
 ➔ *It could give me worth experience **and** help me in my future career*
 ➔ **It could give me worth experience **also** help me in my future career*

As can be seen, the version of example (14) where the expression was substituted only by the adverbial *also* is grammatically incorrect. Although this test indicates that the coordinating conjunction *and* could be used as a substitute of the construction, it cannot be said that it is its head. The whole construction *and also* expresses an additive meaning, and the adverb *also* highlights its expansive property. For these reasons, we considered this construction as one connective in medial position. In the construction *and exactly*, both words indicate a relation of expansion, but a different subtype. *And* is used as to list another argument, but it is the adverb *exactly* that adds the relation of specification:

- (15) *I am twenty years old student of Economic University in Prague **and exactly** I am studying business and law.*

The intended relation in example (15) is specification, but the connective is, again, used wrongly. The second argument in the sentence could be paraphrased as *in specific, I am studying business and law*, in which case the addition of the connective *and* in place of a comma or full stop is superfluous. This example was evaluated as two separate connectives – one medial coordinating *and*, and an initial specifying *exactly*.

There were two cases of multiple connectives which combined different semantic relations. These appeared in Texts 100 and 102:

100	but after all	verbal	expansion + concession	medial
112	and moreover (because)	verbal	expansion + reason	medial

Table 13: Superfluous Use of Connectives

Text number 100 combines the use of the contrastive conjunction *but* with the connective *after all*, which normally expresses concession. However, the context shows that the prepositional phrase is used in a different way:

- (16) *It's been a tough decision for me to choose what would I do in the future during my studies at the grammar school, **but after all** I found myself in leading people, working on projects, meeting deadline and being very enthusiastic.*

In example (16), *after all* is not used to express a meaning of concession, but as an anaphora, where the pronoun *all* refers to a list of activities mentioned by the speaker earlier. The only

connective in this segment is the coordinating *but*. In text number 112, the combination of the connectives *and moreover* is superfluous:

- (17) *I would love to take the opportunity to attend the Barcelona's University of Economics **mainly because** of its beautiful city **and moreover because** of its one of the highest ranking as for the university level of education.*

Both connectives in (17) express the same kind of relation (expansion) and their simultaneous usage is unnecessary, especially since they are followed by yet another connective, *because*.

As can be seen, there were many cases of multiple constructions in the texts that unnecessarily combined the use of two or more connectives. In some cases, this combination was superfluous since the connectives expressed the same kind of relation (e.g. *and moreover*). In other cases, the speakers wrongly combined two distinct relations in one (e.g. *but after all*).

4.2.2 Adverbs - Conjuncts

The second most frequent word class of connectives were adverbs. All adverbs that had a cohesive function were conjuncts¹², but not all conjuncts were realized by adverbs. Below is a table with all the conjuncts used by the B2 speakers, including columns with the total number of occurrences, semantic type, and realization form. The items are ordered according to their frequency in the texts, and the classification of the types of conjuncts is based on Quirk et al. (1985) and Dušková et al. (2006).

CONJUNCT	NUMBER OF OCCURRENCES	TYPE	REALIZATION FORM
So	58	Causal	Adverb
Also	46	Listing	Adverb
Moreover	11	Listing	Adverb
For example	7	Appositive	Prepositional phrase
Therefore	7	Causal	Adverb
Exactly	6	Appositive	Adverb
Especially	5	Appositive	Adverb
First of all	4	Listing	Adverb
Firstly	3	Listing	Adverb
However	3	Contrastive	Adverb

¹² Conjuncts are adverbials with the role of conjoining independent units. They have a relatively detached role as compared with other clause elements, and function as sentence modifiers and means of textual cohesion (Quirk et al., 1985).

In conclusion	3	Summative	Prepositional phrase
Well	3	Transitional	Adverb
Apart from (that)	2	Transitional	Prepositional phrase
Besides	2	Transitional	Prepositional phrase
Nevertheless	2	Contrastive	Adverb
On the other hand/side	2	Contrastive	Prepositional phrase
Secondly	2	Listing	Adverb
Specially	2	Appositive	Adverb
To sum up	2	Summative	Infinitive
After all	1	Summative	Prepositional phrase
At the beginning	1	Listing	Prepositional phrase
Concrete (concretely)	1	Appositive	Adverb
Despite of	1	Concessive	Prepositional phrase
Finally	1	Listing	Adverb
Furthermore	1	Listing	Adverb
In a summary	1	Summative	Prepositional phrase
Last but not least	1	Listing	Adverb
Lastly	1	Listing	Adverb
Overall	1	Summative	Adverb
Than (then)	1	Inferential	Adverb

Table 14: Types of Conjuncts in the Texts by B2 Speakers

The analysis shows that the most frequent conjuncts were *so*, with a total of 58 occurrences, and *also*, with a total of 46. *Moreover* was used 11 times, and the occurrences of the other conjuncts were marginal. The most frequent kinds of relations expressed were (in order from most frequent to less frequent): listing, causal, appositive, summative, contrastive, transitional, inferential, concessive. The expressions *exactly*, *especially* and *specially* were used incorrectly by several speakers in place of *specifically* or *in specific*:

- (18) *I am studying at Economic university in Prague, **especially** I am studying informatic.*
- (19) *I am studying University of Economics in Prague, **exactly** social-economic demography.*

At the beginning, a prepositional phrase which normally functions as temporal adjunct, was also used incorrectly with the meaning of the listing conjunct *First*:

- (20) ***At the beginning*** *I would like to write what I am learning now.*

The adverb *well* was used as a transitional conjunct 3 times and its function was to shift the topic of the discourse. Two of the occurrences appeared within one text:

- (21) ***Well***, thank you for reading and I hope you will give me a positive response.
(22) ***Well***, I have already had work experience. (...) ***Well*** I think this program can give me experiences, new skills and definitely new friendships.

The adverb *now*, used initially as a discourse marker, can have a similar transitional function as the adverb *well* described above (e.g. ***Now***, let us move on to the next topic.). In the texts by B2 speakers, *now* appeared repeatedly at the beginning of the sentence (96 times in total). However, in all these cases, it was only used to express a situational temporal reference, e.g.:

- (23) ***Now*** I am studying at University of Economics in Prague.

A similar function was fulfilled by the adverb *currently*, which was found in the initial position 23 times, e.g.:

- (24) ***Currently*** I am studying on the Faculty of International relations on the University of Economics in Prague.

These two adjuncts were not regarded discourse connectives.

The fact that most adverbials were listing or causal conjuncts can be attributed to the text type, in which the students were required to list their reasons for applying to the Erasmus program. The examples show that the speakers tried to implement some formal expressions to express relations in text, but most of these uses were incorrect. The most frequent conjunct used was *so*, which is more appropriate in spoken and informal language than the written register.

4.2.3 Prepositions and Prepositional Phrases

Out of the 55 occurrences of prepositional phrases, 22 were conjuncts that are listed in the table above. Other prepositional phrases included the expressions *for purpose*, *that*; *as well as*, *in purpose of*, *in case*, *in addition to*, *in order to*, *for the possibility of*, *in connection with*, *for the reason that*, *for this reason*, *because of*, *for*. Out of these expressions, *as well as* and *in addition to* were used as additive connectives, *in case* as a conditional connective, and the rest expressed causal relations. Some of the causal connectives were not used properly. *In purpose of* and *for purpose*, *that* were incorrectly formed expressions, and their meaning was closer to reason than purpose, such as in examples (25) and (26):

(25) *I have decided to write you **in purpose of** studying...*

(26) *I am writing **for purpose, that** I would like to apply for exchange program...*

A frequent prepositional phrase that was used correctly to express purpose was the connective *in order to*. As a connective, *in order to* was found 12 times, and all cases were initial. *For the reason that* and *for this reason* were also examples of secondary connectives, but each had only one occurrence. The connectives *because of* and *for*, which were more frequent in the texts, are discussed below.

When used together with the preposition *of*, *because* forms what is called a compound preposition, *because of*. Like the conjunction *because*, the preposition *because of* expresses causal relations. The difference is that *because* introduces verbal arguments, while *because of* introduces nominal arguments. In the following example, the second argument introduced by the compound preposition is nominal¹³:

(27) *I am interested in the university **because of** its history and reputation.*

Similarly, *for* can work both as a conjunction and preposition. When used as a conjunction, it links verbal arguments in a causal relation:

(28) *I am writing to you **for** I have always wanted to gain much needed experience by studying abroad.*

In most of the occurrences, however, *for* was used as a preposition to introduce nominal arguments:

(29) *I have choosed it **for** its history, profesors that teach at this university and its educational system.*

(30) *This text I write you, because I want study in foreign country (**for** better English knowledge).*

The difference between the two sentences above lies in the directionality of their arguments. In the first example, *for* introduces the argument that provides a reason for the preceding argument ($Arg1 \leftarrow \text{for } Arg2$). In the second example, *for* introduces the purpose of the causal relation ($\text{for } Arg2 \rightarrow Arg1$).

4.2.4 Clauses

Although most of the connectives were one-word, primary expressions, some connectives were realized by entire clauses that contained a subject and predicate. These constructions were in fact more frequent than prepositional phrases, with a total number of 61

¹³ For more information on nominal arguments, see 4.5.2.

occurrences (vs. 55 occurrences for prepositional phrases). The clausal connectives used by B2 speakers were:

<i>A)</i>	<i>B)</i>
<i>The reason is</i>	<i>That is why</i>
<i>The reason why...is</i>	<i>Which is also why</i>
<i>That is because</i>	<i>Which is also the reason why</i>
<i>The reason is</i>	<i>That's the reason</i>
<i>The second reason is</i>	<i>(and) that is part of why</i>
<i>The last reason is</i>	<i>(and) that is the reason why</i>
<i>So this is the reason</i>	<i>That is the main reason why</i>
<i>The main reason why...is</i>	<i>(and) this is the reason why</i>
<i>My main reasons are</i>	
<i>One of the main reasons why...is</i>	
<i>There are some reasons...the first it that</i>	
<i>Another reason is that</i>	
<i>My reason for writing is</i>	

Table 15: List of Secondary Connectives Realized by Clauses

All these clauses express causal relations which could not be simply expressed by the primary conjunction *because*. Semantically, they are all variants of two core secondary connectives:

A) *The reason* (why Arg1) *is* (Arg2)

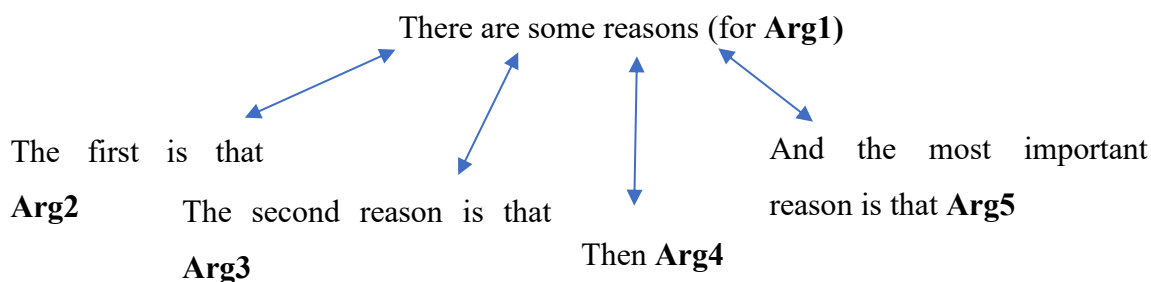
B) *Arg1* (reason) \rightarrow *That is why* \rightarrow *Arg2* (result)

The difference between the two semantic expressions above is their directionality. *The reason is* introduces the reason in the relation, while *That is why* points backwards at a reason stated in the previous argument and introduces the result. The connectives in the left column are all variants of the first expression, and the ones in the right column are variants of the second expression. Most of the connectives involve modification by an adjective, numeral or pronoun: *the first/second/last/main reason is...*; *another reason is that*; *there are some reasons*; *that is the main reason why*. Post-modification by a prepositional phrase is also present: *my reason for writing is*.

Clauses such as *There are some reasons...*, *The/my main reasons are* were used to introduce several semantic arguments at once and joined larger segments of text. In the application letter 82, for instance, a clausal connective joined the sentences in an entire paragraph:

- (31) *There are some reasons, the first is that I want improve my English skills. I think it's best how I can prove my English skill. The second reason is, that I want be able to speak with anybody in the world. Then I want discover some new cultural and business habits. And the most important reason is that after school I would like work abroad and this is the best way, how I can get some experience.*

To explain the cohesive ties brought about by the first connective (*There are some reasons*), perhaps it is best to use a diagram:



Argument 1, which is the resultative part of the causal relation, is in the preceding sentence in the text (Arg1= *I want to try Erasmus student exchange program*). The connective *There are some reasons* introduces a list of arguments that are supposed to explain the wish expressed in the first sentence. The connectives that introduce arguments 2, 3, 4, and 5, are all causal, but they also have an expansive function. Each argument adds new information to the context, a new reason why the speaker should be chosen to attend the Erasmus program. Out of all the texts written by B2 speakers, this is the sample with the largest interconnected segment of text.

4.3 One-Word/Multi-Word, Primary/Secondary Connectives

As defined at the end of chapter 2.3.4.3, connectives can be simple one-word expressions (*and, or, firstly, secondly*) or multi-word constructions (*to sum up, in conclusion, etc.*). The ratio of one-word vs. multi-word connectives in the texts was 823:187. This means that one-word expressions constituted about 81% of the cases, and multi-word expressions only 19%. A graph of this ratio is shown below.

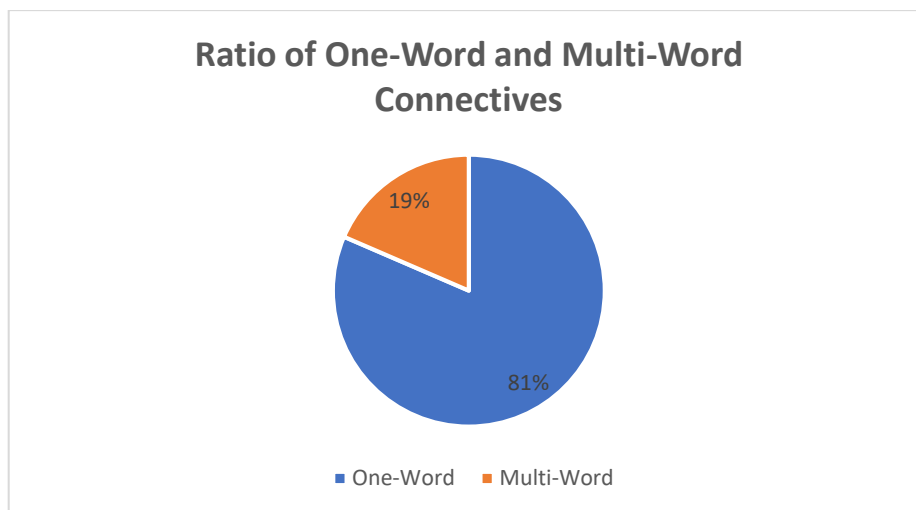


Figure 11: Ratio of One-Word and Multi-Word Connectives

Further, most of the connectives in the texts by B2 speakers were primary connectives – morphologically inflexible, lexically stable grammaticalized expressions that do not normally act as grammatical constituents of a sentence. Most primary connectives were one-word expressions that could be attributed the word classes of conjunctions (*and, because*) or adverbs (*however, also, therefore*), and some primary connectives were realized by prepositional phrases that are grammaticalized as linking elements, or conjuncts (*for example, after all*).¹⁴ All secondary connectives were multi-word expressions and were realized by clauses (e.g. *The reason is...*) or less frequent prepositional phrases (e.g. *for the purpose of*). The respective numbers of primary, secondary, one-word and multi-word expressions are shown in the diagram below.

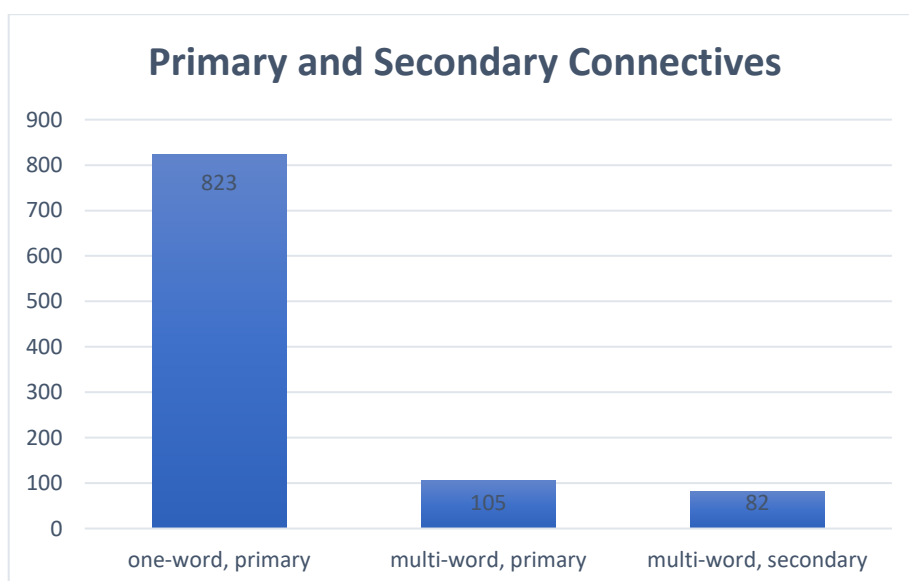


Figure 12: Frequency of Primary and Secondary, One-word or Multi-Word Connectives

¹⁴ For more detail on primary and secondary connectives, see 2.3.5.1.

The diagram illustrates that all one-word expressions (adverbs, conjunctions) were primary connectives. Part of the multi-word expressions were primary connectives (e.g. correlative pairs such as *not only...but also*; conjuncts such as *for example*), and part were secondary connectives, including all clausal connectives. No one-word, secondary connectives were found. Some of the connectives marked as multi-word expressions were redundant uses. The list of these expressions is in the table below.

Text no.	Connective	Type of Arg.	Semantic Type of Relation	Position	One-/multi-word
67	and therefore	verbal	expansion, conjunction	medial	multi
106	and moreover	verbal	expansion, conjunction	medial	multi
112	and moreover because of	nominal	causal, reason	medial	multi
112	and moreover	verbal	expansion, conjunction	medial	multi
116	moreover because	verbal	expansion, conjunction	medial	multi
117	and lastly	verbal	expansion, list	medial	multi

Table 16: List of Redundant Connective Expressions

In all these cases, the use of two connectives instead of one is unnecessary. This is most evident in the combination of the connectives *and moreover*, which both express the same kind of semantic relation, expansion. The combination of the conjunct *therefore* with a coordinating conjunction is also incorrect.

4.4 Position in Sentence

4.4.1 Overview

According to Halliday and Hasan (1976), a word in its cohesive role typically occupies the first position in the sentence, and its meaning extends over the entire sentence (232). Fraser (2009) does not restrict connectives to initial position, saying that they may occur medially or even finally within the second segment of a sequence of discourse. In this analysis, we will differentiate between four different positions of connectives in the sentence: initial, post-initial, medial and final. Since the terms may be misleading in relation to the sentence, let us explain what each position means.

1. Initial

A connective in the initial position occupies the first place in the argument it introduces:

***Because** I'm sick, I can't go to school.*

*I can't go to school, **because** I'm sick.*

The initial position in the first example is clear: *Because* stands at the beginning of the sentence and is part of argument 1. The structure of the sentence is ***Because*** Arg1 – Arg2. In the second example, the fact that the subordinating conjunction *because* stands in the middle of the sentence does not affect the fact that it is only part of argument 2, and appears at the beginning of this argument, hence the term “initial”. The structure of the sentence is Arg1 – ***because*** Arg2.

2. Post-initial

The term “post-initial” means that the connective is part of one of the arguments, but does not appear initially in that argument:

*I can speak Spanish fluently. I can **also** speak some French.*

Here, each sentence represents one argument. The adverb *also* is part of the second argument (*I can speak some French*), but occupies a position other than initial/final within this argument.

3. Medial

A connective in the medial position is not part of either of the arguments that are in a semantic relation – it stands independently between them. Connectives that typically appear in the medial position are coordinating conjunctions (*and*, *but*, *or*):

*I'm sick **and** I'm feeling terrible.*

The structure of the example above is *Arg1* – **and** – *Arg2*. The connective is not part of argument 1 nor argument 2, and it occupies the medial position between them. A connective is said to have a medial position even if its scope stretches beyond the sentence, but it is not part of either of the arguments:

I'm sick. And I'm feeling terrible.

The structure of this utterance is the same as the previous example, even though there is a full stop between the arguments: *Arg1* – **And** – *Arg2*.

4. Final

Some connectives can even occupy a final position within the argument they are part of:

We don't have to go. I will go, nevertheless. (Fraser, 1999: 938)

Most of the connectives in the texts written by B2 speakers were initial. Below is a table with an overview of the different positions of connectives in the texts, and a figure illustrating the phenomena.

Position in sentence	Total Number of Connectives
Initial	528
Post-initial	24
Medial	454
Final	6

Table 17: Position of Connectives in Sentence: Initial, Post-Initial, Medial, Final

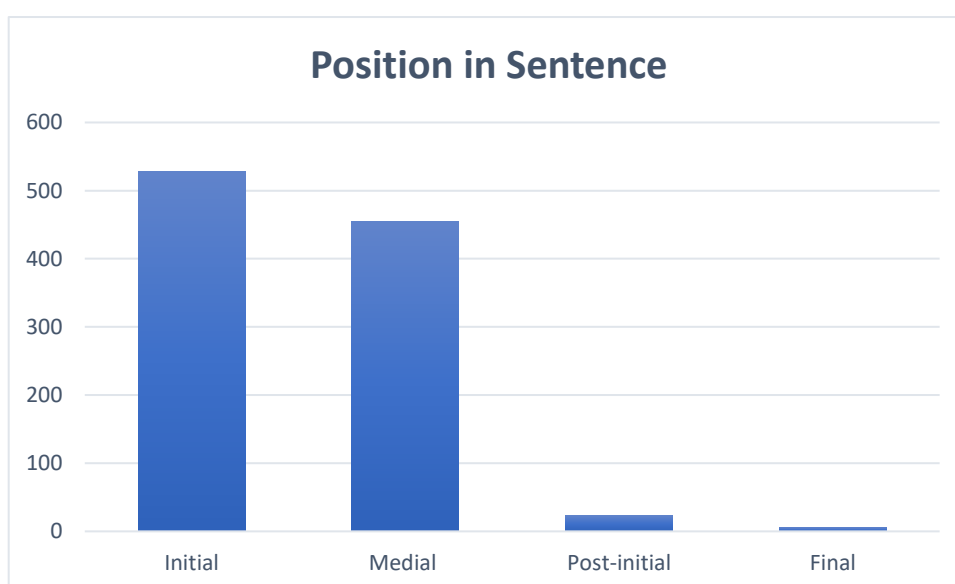


Figure 13: Position of Connectives in Sentence: Initial, Post-Initial, Medial, Final

The graph above shows that the most frequent position was initial, which constituted more than half of all examples (56%). The second most frequent position was medial (41%), then post-

initial (2%) and final (1%). Most of the connectives had a preferred position in the sentence, which could be due to some of their inherent properties. Subordinating conjunctions (*because, if, since, after, although, when, while, before*), whose function is to link subordinating clauses to main clauses, were found almost exclusively initially:

- (32) *My wanted specialization is in Marketing sector, **because** I'm already working in Prague radio station in this department. (Arg1 – **because** Arg2)*

The coordinating *and, but, and or* were found only medially:

- (33) *I'm a Vietnamese **but** my nanny was Czech. (Arg1 – **but** – Arg2)*

Some connectives were not tied to a certain position and oscillated between two uses. Conjuncts, which have the function of joining independent units, appeared mostly at the beginning of the sentence, but sometimes in post-initial, or even final positions (*too* and *as well*, see 4.4.3). The adverb *also* appeared mostly post-initially, but sometimes initially. Examples of both uses are in the table below.

Position	Text no.	Example
Initial	34	<i>Also, when I was at the high school, I was part of a project</i>
Post-initial	10	<i>I also love Paris and France in general</i>
	14	<i>I also have finished English course at level B2</i>
	40	<i>I have also taken countless seminars on legal problems and their solutions.</i>
	118	<i>I am a native Czech speaker and I also speak English and a little bit German.</i>

Table 18: Position of the Connective "also" in the Texts by B2 Speakers

There were some errors in the use of *also* post-initially. Three speakers placed the connective before the auxiliary instead of after, such as the author of text number 14:

- (34) *I **also** have finished English course at level B2 (instead of I have **also** finished English course at level B2).*

These cases were nevertheless considered and counted as post-initial. Notice that the sentence in the example from text 118 contains both a medial coordinating conjunction (*and*) and a post-initial adverb (*also*). This chunk, *and I also*, was found 6 times in the texts.

To sum up, most connectives were found initially, i.e. at the beginning of the argument they introduced. A significant number of connectives was found medially, which is due to the frequent use of the coordinating *and*. Most connectives also had a preferred position in the sentence, though some connectives, namely conjuncts, appeared in two different positions

within the argument. There were no cases of connectives that could operate both as medial and as initial/post-initial/final.

4.4.2 Position and Scope (Intra-/Inter-Sentential)

In relation to the position of connective it is interesting to consider their scope. In this chapter, we will analyze whether the connectives were inter-sentential or intra-sentential (i.e. whether they expressed a relation between two arguments within one sentence, or whether their meaning extended to a larger utterance, across the border of the sentence in which they appeared). We will also discuss the directionality of the connectives (i.e. the order of the arguments they joined).

In most cases, the scope of the connective was only over the sentence in which it appeared. In total, there were 856 intra-sentential connectives and only 155 inter-sentential connectives. The ratio between inter- and intra-sentential connectives is shown in the diagram below:

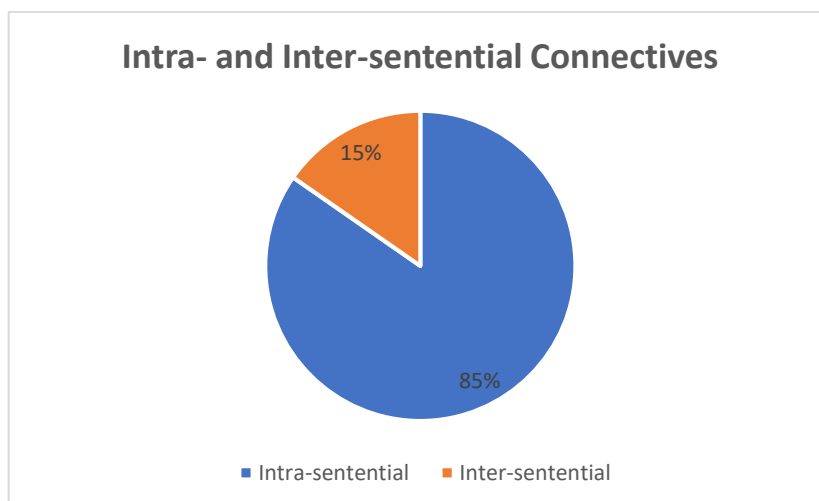


Table 19: Ratio of Inter- vs Intra-Sentential Connectives

A typical example of a connective that joined two arguments within one sentence is the following:

- (35) *I am interested in your institution **because** it is the best in what I want, in management.*

In example (35), the main clause contains argument one (*I am interested in your institution*) and the subordinate clause introduced by the subordinating conjunction *because* contains argument two (*it is the best in what I want*). There is a causal relation between the two arguments: *Arg1 – because Arg2 (reason)*. The sentence that follows in the text, on the other hand, begins with a connective:

(36) ***The second reason is that*** *Wonderland is the most beautiful country in the world.*

Here, the connective in the initial position extends over the border of the sentence. It introduces another argument (Arg2) that is related to the first argument contained in the previous sentence (Arg1) and forms another causal relation of reason: ***the second reason is that*** – reason (Arg2). The connective, realized by the construction *subject – copula*, is secondary. The subject complement of the whole construction is expressed by the nominal content clause *Wonderland is the most beautiful country in the world*. Such initial inter-sentential connectives are, however, not as frequent in the texts written by B2 speakers. In fact, many of the connectives that occupy an initial position have a restricted scope. In 18 cases, the initial connectives are intra-sentential, meaning that they express a relation only within the sentence. Such is the case in conditional clauses introduced by the subordinating conjunction *if*:

Text no.	Example
4	If <u>I can get a possibility</u> to go to Oxford University it surely will be good experience.
6	If <u>you will take me</u> , it will be very good valued experience for my future job and also for my life.
13	If <u>I was chosen for the program</u> I would also be interested in a part time job
148	If <u>you are interested in my skills</u> , experiences, etc., then you can trust, that (...)
149	If <u>you want to make some international contract</u> you will need to be able to communicate with your partners
152	If <u>you are interested</u> you can contact me on (...)

Table 20: Initial Intra-Sentential Use of "If"

In all these examples, the order of the arguments in the sentence is inverted. The subordinating conjunction placed at the beginning of the sentence introduces a subordinate adverbial clause of condition, and the second clause (Arg1) is the main clause. A similar case occurs when the relation expressed is concessive:

(37) ***Although*** *I rely not only on academic education*, *I would be extremely happy to be accepted;*

and with some temporal references, such as:

(38) ***Before*** *I started my studying at university*, *I attended gymnasium.*

Even though the connectives in these examples express intra-sentential relations, the sentences are not entirely taken out of context and they remain cohesive. The semantic ties are provided by other means of cohesion, namely by reference and lexical cohesion (reiteration). In the first

example from text number 11, the lexical item *academic education* is a superordinate term for other items mentioned in the previous linguistic context. In the second example, cohesion is achieved by the possessive pronoun *my* in the noun phrase *my studying at university*, which is an example of personal reference.

There were 154 cases where it was the connective that formed cohesive ties. In most of these examples, the connective was realized by an initial conjunct (e.g. *however, therefore, to sum up*). Consider the following sentences, where the connective relates an argument to previous linguistic context:

(39) *CBS has a great reputation as one of the best business schools in EU and the world. **So** it will be a great contribution to my future success.*

(40) ***Besides**, I have been studying German since my first year in secondary school.*

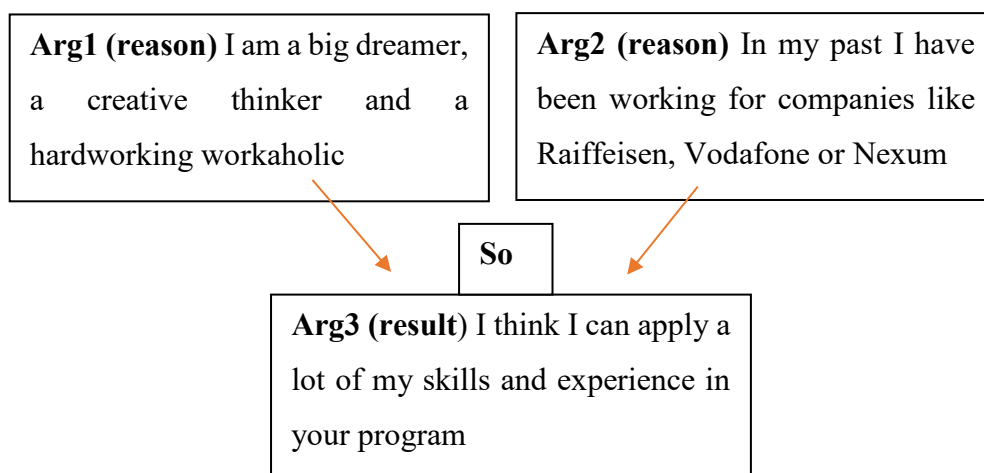
Notice that, although both connectives in the examples above are inter-sentential, their scope is different. In the first example, the causal conjunct *So*, which is the less formal variant of the conjunct *Therefore*, relates the argument to the statement in the previous sentence. In fact, the full stop in the first sentence could be easily replaced by a comma and the cohesion of the utterance would not be violated. The cohesive tie created here is: *Arg2 – So – Arg2 (reason)*. In the second example, on the other hand, the transitional conjunct *besides* relates to several preceding sentences at once:

*Before I started my studying at university, I attended gymnasium in my hometown. I've been studying English for eleven years now. **Besides**, I have been studying German since my first year in secondary school.*

So may also have a greater scope as an inter-sentential connective, as in the following example from text 69:

(41) *I am a big dreamer, a creative thinker and a hardworking workaholic. In my past I have been working for companies like Raiffeisen, Vodafone or Nexum. **So** I think I can apply a lot of my skills and experience in your program.*

In this case, the semantic meaning of the connective is not only causal (result), but also summative. The speaker first lists his/her qualities and achievements that lead up to the statement made in argument 3. For a better idea, the following diagram illustrates the semantic relations between the three sentences:



As is shown in the diagram, the conclusion/result introduced by the connective *So* is based both on argument 1 and argument 2; the connective extends beyond the sentence it introduces.

4.4.3 Final Connectives

Interestingly, there were several cases where the connective was placed at the very end of the sentence, which is not a common position. Five occurrences were of the connective *as well*, and one of the connective *too*. In all cases, the connective expressed an additive relation:

- (42) *I have been studying Spanish at my former high school and also obtained certificate at L1 level. This year I passed my CAE exam **as well**.*
- (43) *I have worked in a small corporate in my country, I my work was about planning strategy. I worked on my own projects, **too**.*

Although the argument precedes the connective in both examples above, the role of the expressions *as well* and *too* is cohesive, since they link the entire sentence to the preceding context. The sentences could be paraphrased using other non-final connectives, such as *also* or *further*:

*I have **also** passed my CAE exam this year. **Further**, I have passed my CAE exam this year.*

4.5 Discourse Arguments

Now that we have defined and classified the classes and types of connectives, we will focus on the analysis of discourse arguments and relations. Discourse connectives in the strictest sense are defined as linguistic expressions that join two verbal arguments, i.e. segments of text that contain a finite verb (Rysová, 2015: 64). However, when taken from a semantic perspective, discourse arguments do not always have to be verbal, provided that the meaning of the utterance and the semantic relation between the two arguments is maintained. In the analyzed texts, the vast majority of arguments were verbal (over 97%), but there were some

cases of nominal arguments too (26 occurrences). 84 arguments were realized by infinitives, but since they all involved ellipsis of a finite verb, they were grouped under verbal arguments.

4.5.1 Verbal Arguments

As verbal arguments we considered all arguments that contained a finite verb, e.g.:

- (44) *I am contacting you **because** I am interested in participating at the Erasmus student exchange program.*

In example (44), argument 1 contains a verb in the present continuous (*I am contacting*) and argument 2, introduced by the causal connective *because*, contains a copular verb (*I am interested*). In the example below, the coordination of verbal arguments involves ellipsis:

- (45) *I would practice my English a lot **and** learn new stuff*

In example (45), the subject and modal auxiliary are elided in the second argument (*~~I would~~ learn new stuff*). The meaning of the utterance is clear, and the elided elements are recoverable from the context. In a written register, repeating such elements would only lead to redundancy

***I would** practice my English a lot and **I would** learn new stuff.*

Repetition is a typical feature of spoken language, but in the written register, clarity and accuracy is preferred.

There was one case where an argument seemed to be realized by a past participle of a verb:

- (46) *I have always wanted to live in London **and** gotten to know English culture better*

However, on a closer look, the example contains a grammatical mistake. The past participle is misused instead of an infinitive. The grammatically correct sentence would be:

*I have always wanted to live in London **and** to get to know the English culture better*

Assuming this is the meaning intended by the speaker, the coordinating conjunction fulfills a cohesive function here. The semantic meaning is additive, and the order of the clauses cannot be switched without violating the cohesion of the sentence. Although the following sentence with inverted clauses would be grammatically correct, it would lose its logical development:

I have always wanted to get to know the English culture better **and to live in London.*

Infinitives are, of course, not finite verbs, but since the use of infinitives in coordinated structures always involved ellipsis of a finite verb that was present in the immediate linguistic context, we considered arguments that were realized by infinitives as verbal. Consider the following example of verbal ellipsis:

- (47) *This would help me continue on my studies **and** to improve my language skills.*

Here, the second argument is realized by an infinitive (*to improve*). Due to the coordination of clause elements, the finite verb form that is present in the first argument (*would help continue*) is

elided. There were 84 cases where an argument was realized by an infinitive, which constitutes about 8,3% of all examples.

4.5.2 Nominal Arguments

There were some examples in which one of the arguments in a semantic relation was nominal. This means that the argument did not contain a finite verb. In the following example (48), there is a causal relation between argument 1 and argument 2, which is expressed by the preposition *because of*:

- (48) *I had to quit them **because of** health problems.*

Here, the semantic relation between the two arguments (causal, reason) is maintained, even though the argument introduced by the connective is nominalized. The example could be paraphrased using a finite verb:

*I had to quit them **because of** health problems → I had to quit them **because** I had health problems*

When both arguments are verbal, the semantic relation between them is more evident. However, the fact that the second argument is nominalized does not change the semantics of the sentence. As explained in 4.2.3, *for* as a preposition can also introduce nominal arguments:

- (49) *I have choosed it **for** its history, profesors that teach at this university and its educational system.*

However, most of the cases where an argument was nominal involved grammatical mistakes on the part of the speakers. Consider for example the use of the adverb *exactly*:

- (50) *I am studying University of Economics in Prague, **exactly** social-economic demography.*

- (51) *I study at University of Economics in Prague, **exactly** business economic and management*

Although the adverb *exactly*, which is still considered a connective due to its cohesive role (*Arg1 – **exactly** Arg2*), introduces a noun phrase in both examples, the underlying relation between the two arguments is that of specification. *Exactly* is used with the meaning of *specifically* or *in specific*. The first argument, which is verbal, is a general statement about the speaker. The second nominal argument restates and at the same time specifies what was said in the first argument. The finite verb in both clauses is implied (***exactly**, I study...*).

Errors also appeared in the causal use of *since*:

- (52) ***Since** more experience, meet with new interesting people and other cultures in the world...*

The semantic relation that can be deduced from the preceding sentence (*I want to improve my skills in English and another language*) is that of purpose (*so I gain more experience...*). The sentence is grammatically incorrect, but there is an implicit discourse relation between the utterance. In other examples, *since* was used merely as a preposition:

112	Since <u>the grammar school</u> I have been always taking a part in interesting projects
123	I have been studying the French language since <u>the 1st grade at school</u>
127	Since <u>2011</u> I've been working as a freelancer in multimedia business

Table 21: "Since" used as Preposition

In these cases, *since* is part of a prepositional phrase that functions as a temporal adverbial within the clause but does not have a cohesive role between two arguments.

As can be seen, the speakers had a certain difficulty in the use of connectives together with nominalized arguments. The common mistake was the inappropriate use of ellipsis in segments where the elided element was not clearly recoverable from the utterance, and some expressions were used to express a different semantic relation than their core meaning.

4.6 Semantic and Pragmatic Types of Relations

4.6.1 Overview

Each example was analyzed individually to determine what kind of semantic relation the connective expressed. The chosen model of categorization was the hierarchy of semantic relations in the PDTB (2007, see 2.2.3), which divides relations (*senses*) into temporal, contingency, comparison and expansion. The total number of connectives that expressed these four general types is shown in the table below.

Semantic Relation	Number of Occurrences
Expansion	600
Contingency	336
Comparison	55
Temporal	17

Table 22: Semantic Types of Relations: Expansion, Contingency, Comparison, Temporal

The most frequent type of relation was that of expansion, the second most frequent relation was that of contingency (which includes the subtypes of reason and result). Further, 55 connectives expressed relations of comparison, and only 17 connectives were temporal. The diagram below shows the numbers of the main types of relations in percentages, demonstrating the prevalence of expansive relations.

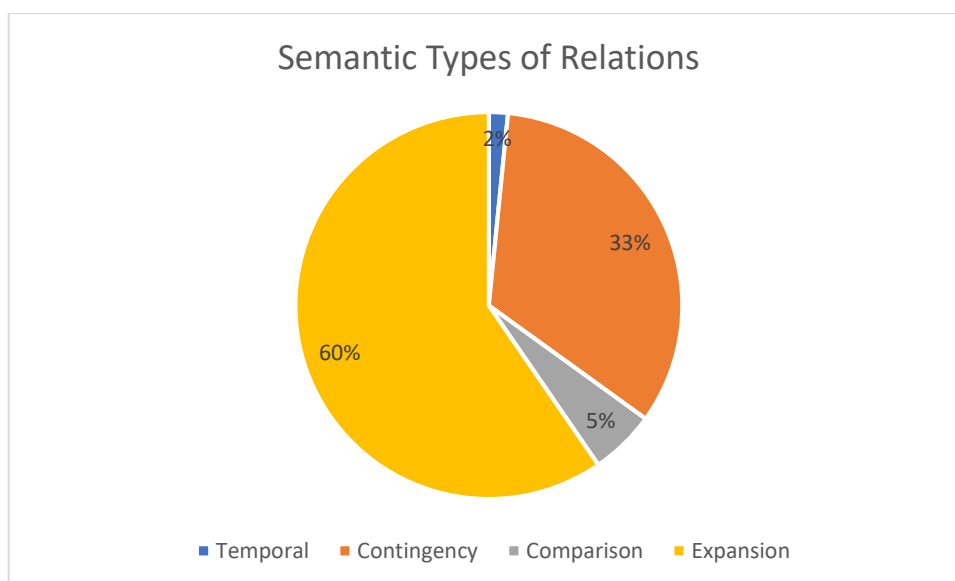


Figure 14: Ratio of Semantic Types of Relations

The fact that 60% of the relations expressed were that of expansion corresponds with the high incidence of the additive conjunction *and* (with 372 occurrences in total). For the B2 speakers, coordination of shorter structures was the most frequent way to add arguments. The frequency of relations of contingency (33%) can be attributed to the character of the text type (the application letter), in which the speakers were required to provide the reasons for applying for the Erasmus exchange program. Temporal relations and relations of comparison were marginal.

Each category (temporal, contingency, comparison, expansion) was further divided into several subcategories that specified the type of relation expressed by the connective. Not all subcategories as described by the PDTB were found in the texts by B2 speakers. The following table contains an overview of the subtypes of semantic relations that were present in the texts, their frequency in the texts, and the connectives that were used to express these kinds of relations.

Semantic Relation	Subtype	Number of Occurrences	Connectives
Temporal	Asynchronous, succession	10	<i>Then, after, since</i>
Temporal	Asynchronous, precedence	3	<i>Before</i>
Temporal	Synchronous	1	<i>When</i>
Contingency	Cause, reason	226	<i>Because, because of, since, for; The reason (why) is¹⁵</i>
Contingency	Cause, result	72	<i>So, therefore</i>
Contingency	Condition, hypothetical	19	<i>If, especially if, when, in case</i>
Contingency	Cause, purpose	17	<i>In order to, for, for the possibility of, the purpose of which is; That is why¹⁶</i>
Contingency	Condition, general	2	<i>If</i>
Comparison	Contrast	50	<i>But, while, on the other side, on the other hand</i>
Comparison	Concession	5	<i>Despite of, although, nevertheless, but after all</i>
Expansion	Conjunction	421	<i>And, also, and also, not only...but also, not just...but also, in addition, in addition to, moreover, and moreover, besides, furthermore, apart from, as well as</i>
Expansion	List	133	<i>And, also, then, at the beginning, firstly, secondly, (and) lastly, last but not least</i>
Expansion	Restatement, specification	17	<i>Exactly, concretely, especially, specially, and, and exactly, to be more exact</i>
Expansion	Instantiation	13	<i>For example, the first thing is, first of all, secondly</i>
Expansion	Restatement, generalization	10	<i>In conclusion, to sum up, in a summary, overall, so, well, there are all the reasons</i>
Expansion	Alternative	6	<i>Or, either...or</i>

Table 23: Types of Semantic Relations and Corresponding Connectives in the B2 Texts

¹⁵ The variants of the secondary connective *The reason is* are found in 4.2.4.

¹⁶ The variants of the secondary connective *That is why* are found in 4.2.4.

As can be seen, most of the relations above could be expressed by more than one connective. Only the categories of synchronous temporal relations, relations of precedence, and general condition, were expressed by a single type of expression (*before*, *when*, and *if*, respectively). The table also indicates that some connectives could be used to express different kinds of relations. Below is a table that lists some of the most frequent connectives and the semantic relations they expressed in the texts.

Connective	Relation Expressed
And	<i>Expansion, conjunction; expansion, list; expansion, specification</i>
But	<i>Comparison, contrast</i>
Or	<i>Expansion, alternative</i>
For	<i>Cause, reason; cause, purpose</i>
Because	<i>Cause, reason</i>
Because of	<i>Cause, reason</i>
If	<i>Contingency, condition</i>
Since	<i>Cause, reason; temporal, succession</i>
Also	<i>Expansion, conjunction; expansion, list</i>
So	<i>Cause, result</i>
Therefore	<i>Cause, result</i>
For example	<i>Expansion, instantiation</i>
Exactly	<i>Expansion, specification</i>

Table 24: Overview of Most Frequent Connectives and Semantic Relations They Express

The table shows that the relation between the connectives and the discourse relations they expressed was mostly one-to-one. This was true for the connectives *but*, *or*, *because*, *because of*, *if*, *so*, *therefore*, *for example*, and *exactly*. Some connectives, on the other hand, could be used to express different kinds of relations, depending on the specific case. These were *also*, *since*, *for* and *and*. The coordinating conjunction *and* was most poly-functional of all, since it was used to express 3 different types of relations.

4.6.2 Semantics of “And”

In the texts by B2 speakers, *and* was used to express different subtypes of expansion. The PDTB (Prasad et al., 2007) distinguishes between two additive (expansive) uses of a connective – *conjunction* and *list*. The description of these two relations is in the table below.

Subtype of Expansion	Definition	Examples provided by the PDTB
Conjunction	The situation described in Arg2 provides additional, discourse new, information that is related to the situation described in Arg1	<i>Also, in addition, additionally, further</i>
List	Arg1 and Arg2 are members of a list, defined in the prior discourse; the situations specified in Arg1 and Arg2 do not have to be directly related	<i>And</i>

Table 25: Subtypes of Expansion (Conjunction, List) as Defined by the PDTB (2007)

The difference between the two categories is that, in a conjunctive relation, the information described in argument 2 somehow relates to the information provided in argument 1, while in *list*, the two arguments do not have to be directly related. In the texts by B2 speakers, most uses of *and* were conjunctive (69%). The second most frequent relation was *list* (29%), then *specification* (2%). The table below shows the frequencies of the individual relations expressed by *and*, with examples in each category.

Semantic Relation	Frequency	Examples
Expansion, conjunction	256	159: I know, that your university is one of the best in the world and a lot of students would like to have a chance to study here. 160: I would like to once again point out, that I am very keen on studying language and that for me there is no target line to stop studying them.
Expansion, list	107	143: My name is XY and I would like to apply for the Erasmus student exchange program at some university in England. 155: Now I am studying in Czech Republic in University of Economics and I want to improve my skills of English language.
Expansion, specification	7	144: I study University of Economy in Prague and my specialization is information technology.

Table 26: Semantic Relations Expressed by “and” in the B2 Texts

The examples in the table above demonstrate the differences between the relations of conjunction and list. In the sentences where *and* is conjunctive, the arguments are related not only semantically, but also syntactically. In both cases, *and* joins subordinate clauses that function as the multiple direct object of the whole sentence. In the uses of list, *and* links clauses that are not semantically related and could appear as separate sentences.

There were many instances (in total 32) where *and* had an additive function within a more complex syntactic structure. Consider the following example:

- (53) *I would like to study at your university because I am keen on French **and** I want to improve my French language skills.*

Here, the connective *and* could be substituted by *and because* to make the causal relation between the arguments clear. However, the function of *and* within this relation is conjunctive – it expresses an additive relation between two adverbial subordinate clauses of reason. For a visual representation, the syntactic and semantic relations in the example (53) are depicted below:

Arg1	because	Arg2		
I would like to study at your university		I am keen on French (Arg1)	and	I want to improve my French language skills (Arg2)

To conclude, the connective *and* was used by the B2 speakers to express three different types of semantic relations, but all within the category of expansion. Its function was either to add an argument related to the previous sentence, an argument that was only part of a list, or an argument that specified previous context.

5 CONCLUSION

The aim of this thesis was to analyze the use of discourse connectives in formal texts written by non-native speakers of English, specifically of speakers whose language proficiency corresponds to the level B2 in the Common European Framework of Reference. A detailed analysis was carried out of the frequency of discourse connectives, their types, use, and of the semantic relations they expressed in the texts. The results indicated some interesting trends, which will be presented in this chapter.

In total, 1011 connectives were found in the 161 texts written by B2 speakers of English, which gives an average value of 6,28 connectives per text. The fact that not all texts were of the same length was considered and the ratios of connectives per sentence and sentences per connectives were calculated. On average, there were 0,63 connectives per sentence, and 1,8 sentences per one connective. The minimum number of connectives per text was 1 and the maximum was 17, which gives a range of 16 connectives. The median value, or the most frequent number of connectives per text, was 4. For further insight into the use of connectives by B2 speakers, we looked at them from several different aspects, namely: morphological/syntactic categories, types of connectives (one-word/multi-word, primary/secondary, modified/parallel/conjoined), their position in the sentence, the types of arguments (verbal/nominal), and the semantic types of relations that the connectives expressed. The outcomes of these individual sections will be presented as follows.

We worked with the hypothesis that most connectives would be simple one-word expressions from the classes of conjunctions, adverbs and prepositional phrases. The results showed that conjunctions were indeed the most frequent morphological class, with a total of 675 occurrences. The second most frequent class were adverbs, with a total of 208. In 61 cases, the connectives were realized by clauses, and in 55 cases by prepositional phrases. These values form relative percentages of 68% (conjunctions), 21% (adverbs), 6% (clauses) and 5% (prepositional phrases). The ratio of one-word vs. multi-word connectives in the texts was 823:187, meaning that over 80% of the connectives were single-word expressions. Most of the connectives were primary; all secondary connectives were multi-word expressions realized by clauses (e.g. *The reason is...*) or prepositional phrases (*for the purpose of*). Some multi-word expressions were primary connectives (e.g. correlative pairs like *not only...but also*, conjuncts such as *for example*).

Generally, the speakers used a variety of connectives in their texts, but when considered proportionally, only a few expressions made up most of the instances. These were the

coordinating conjunction *and*, the subordinating conjunction *because*, and two adverbs, *so* and *also*. These connectives were used across the texts by many different speakers. The rest of the connectives, which included other conjunctions, adverbs, prepositional phrases, and clauses, were used unsystematically by specific speakers. The same trend was found in the studies by Chen (2006) and Lahuerta Martinez (2004), who both concluded that students of English relied on a limited set of connectives in their writing.

More than 2/3 of the conjunctions used by B2 were coordinating, which suggests the prevalence of paratactic relations. The coordinating conjunction *and* occurred 372 times in total¹⁷, which constitutes roughly 37% of all connectives. This also constitutes a strong majority within coordinating conjunctions, since the contrastive *but* appeared 55 times, the alternative *or* only 4 times, and *for* also only 4 times. These values mirror the results of the study by Aysu (2017), which also indicated a very high frequency of the coordinating *and* (98 instances out of 180, which is about 54.4% of the occurrences). In the texts, the coordinating *and* had purely expansive functions. Following the classification of semantic types of relations by the PDTB (Prasad et al., 2007), we found that *and* expressed three subtypes of expansive relations: conjunction, list, and specification. This made the coordinating conjunction *and* the most poly-functional connective of all, since most other connectives were used to express one or a maximum of two types of relations. *And* was used by the B2 speakers either to add an argument related to the previous sentence (conjunction), to enumerate an argument at part of a list (list), or to specify information provided in the previous argument (specification). The most frequent use of *and* was conjunctive, which constituted 69% of the occurrences. The high frequency of the coordinating conjunction *and* suggests the speakers' inability to use more complex structures to express additive relations. There were only 11 instances of the listing conjunct *moreover*, and the listing conjunct *also* (which is also used to express relations of expansion, conjunction) was often used together with the coordinating *and*. Some speakers avoided the use of the coordinating conjunction by chaining sequences of shorter sentences, using asyndetic coordinating or other means of cohesion such as reference or lexical cohesion.

There was a general overuse of connectives by the B2 speakers. The coordinating *and* was sometimes used to connect segments of discourse that were not directly related to each other. It was also used in combination with other connectives that either expressed a different kind of relation (e.g. *and because*), or even the same one (e.g. *and also*, *and moreover*). The speakers formed other combinations of connectives as well, e.g. *moreover because*, *but after*

¹⁷ The coordinating *and* in fact appeared 523 times in total, but in 151 examples, its use was only structural. For more detail on the structural use of *and*, see 2.3.2.

all, etc. Such multiple constructions were unnecessary and redundant, and follow the trends described in the studies by Chen (2006), Shaw (2009) and Tomešová (2017), which all confirmed a strong overuse of connectives by non-native speakers of English. As an example, consider text number 88. All connectives in the text are highlighted:

There is a total number of 12 connectives in this text, which constitutes a density of 1,2

88
Dear Sir or Madam
I am writing you in order to apply for the Erasmus student exchange program which I would like to attend the following semester.
Firstly , I would like to introduce myself. Currently, I am a first year student of the University of Economics in Prague. My major is diplomacy, therefore I am interested in studying the same domain on the partner university as well .
Last year I graduated from high school and received state certificate in English level B1. Moreover , I also study French and passed my final exams on B2 level. During the school year of 2013/14 I was a participant in an Exchange program and spend one year in the USA where I lived in a host family and gained language skills as well as useful experience with new culture and people.
In general I am an open person interested in exploring new opportunities and the idea of spending a whole semester in a foreign country excites me and also opens a whole lot of doors for my future life.
Attending this program would only bring benefits to my career as a student and future diplomat. It would help me to learn how to communicate and also take part in international projects and affaires.
I am looking forward to hearing from you,
Yours sincerely XY

connectives per sentence. There are several cases where the use of a connective is superfluous. For example, in the sentence “My major is diplomacy, therefore I am interested in studying the same domain on the partner university as well” the connective *as well* only repeats what has already been expressed by the modifier *same*. The adverb *moreover* is combined with the adverb *also*, which expresses the same relation of conjunction. The coordinating conjunction *and* is used 6 times, which is half of the total amount of connectives.

Further, many of the connectives were used incorrectly. Most of these mistakes were made in the use of more complex types of expressions such as prepositional phrases, conjuncts and clauses, and they regarded the position of the connectives in the sentence, the semantic relations expressed, and the appropriateness of the expressions in the formal register. Some connectives (e.g. *and*, *because*) that are normally placed in between clauses were placed at the beginning of the sentence. Other connectives (e.g. *exactly*, *especially*) were used to express the wrong type of semantic relation. The adverb *so*, which is more proper to informal language and

the spoken register, was used frequently by the speakers. There were also some ungrammatical uses of connectives, e.g. *concrete* (instead of *concretely*); *in purpose of*; *in purpose*, *that*, etc. This shows that the speakers were not in good command of more elaborate structures and should be instructed on the use of connectives to express different kinds of relations.

In terms of semantic relations, the most frequent type of relation was that of expansion (60%), the second most frequent relation was contingency (33%), which was followed by relations of comparison, and lastly temporal relations. Most of the semantic relations could be expressed by more than one connective. Most connectives, on the other hand, expressed only one or two semantic relations (except for *and*, which was explained before). The high frequency of expansive relations corresponds with the frequent use of the additive *and* (372 occurrences) and *also* (46 occurrences). The frequency of the category of contingency, which includes the subtypes of reason and result (purpose), corresponds with the frequent use of the causal connectives *because* (151 occurrences) and *so* (58 occurrences). The high incidence of causal connectives can be attributed to the text type, where the subjects were asked to write an application letter for the Erasmus program, which required the students to introduce themselves and provide their reasons for applying. Relations of contingency were expressed using not only primary connectives such as *because* or *therefore*, but also more complex secondary structures, such as *The reason is*, *That is why*, etc. In fact, all secondary clausal connectives expressed causal relations. This is due to the fact that the conjunction *because* does not allow the formation of more complex syntactic structures. The text type also influenced the types of conjuncts used. The most frequent types of conjuncts were (from most frequent to less frequent) listing, causal, appositive, summative, contrastive, transitional, inferential, and concessive.

The preferred position of connectives in the sentence was initial, a tendency that was found by Shaw (2009) and Tomešová (2017) for both native and non-native speakers of English. More than half of the connectives in the texts by B2 speakers were initial (56%); 41% were medial, and some connectives appeared in post-initial (*also*) and even final positions (*as well*, *too*). The large percentage of the medial position was due to the high frequency of the coordinating *and*. The other coordinating conjunctions were also found exclusively in the medial position. Most of the connectives had a preferred position in the sentence, though some conjuncts (*also*) appeared in different positions within the argument they were part of (post-initial, initial). The arguments that the connectives joined were mostly verbal (a vast majority of 97%). In 82 cases, the argument was realized by an infinitive, and in 24 cases, the argument introduced by the connective was nominal.

The meaning of most the connectives extended only over the sentence in which they appeared, which is a trend that was also found by Tomešová (2017) in the study of result/inference discourse connectives in academic texts. The scope of all post-initial and medial was intra-sentential (except for cases where the connective was misused, e.g. the use of *and* or *because* at the beginning of the sentence). Many of the initial connectives also had a restricted scope. The initial connectives that had an inter-sentential scope were listing, causal, summative and inferential conjuncts such as *first*, *second*, *therefore*, *in conclusion*, *besides*, etc. The greatest scope was found of the connective *so* and some secondary expressions such as *There are some reasons...the first is...the second is*, which were used to connect larger utterances. All conditional, concessive and temporal conjuncts were intra-sentential.

To conclude, the results agreed with our hypothesis in some ways and differed in others. The speakers used a variety of connectives, but proportionally, most occurrences were of a limited number of connectives, which included the coordinating conjunction *and*, subordinating conjunction *because*, and the adverbs *so* and *also*. Most connectives were then one-word primary expressions realized by conjunctions, adverbials or prepositional phrases. However, a significant number of instances was also realized by clauses. The connectives realized by clauses were all modifications of two basic structures, *The reason is* and *That is why*, and all expressed relations of contingency. We assumed that most connectives would appear medially, which was not confirmed – most connectives appeared initially. The scope of most connectives was immediate, i.e. their meaning extended only over the sentence which they were part of. The main semantic type of relation expressed by the connectives was expansive (additive), which is due to the high occurrence of the coordinating *and*. The second most frequent relation was that of cause, which can be attributed to the text type and the frequency of the subordinating conjunction *because* and of clausal connectives. We also assumed that the frequency of connectives per letter would not be too high. The results indicated the contrary – a general overuse of connectives by B2 speakers, including the use of the coordinating *and* to connect segments of discourse that were not directly related to each other, or the formation of superfluous multiple constructions such as *and moreover because*. Errors were found in the use of both primary and secondary connectives in terms of their position, semantics and stylistics.

Although the results of this thesis did not agree in all points with our hypothesis, and many errors and unclarities were found in the texts by B2 speakers (which regarded not only the use of connectives), the outcome meets the objectives that we defined in the first place. The current descriptive scales in the CEFR do not provide much information about the specific cohesive devices used by non-native speakers of English. The results of this study can serve to

specify the level B2 in the CEFR in terms of writing and text coherence. This could be applied to the assessment of texts written by non-native speakers of English. The outcomes could also be used in language instruction as a guidance in the areas that the speakers should focus on to perfect their writing. Of course, since all analyzed texts belonged to a certain text type, the results cannot be considered entirely objective. Further research would need to be done to draw general conclusions about the use of connectives by B2 speakers of English in written production. Nevertheless, the thesis fulfilled its aims and can be used as a basis for further analysis.

References

- Al-Saif, A. and K. Markert** (2010) 'The Leeds Arabic Discourse Treebank: Annotating Discourse Connectives for Arabic'. Leeds: University of Leeds, School of Computing.
- Ariel, M.** (1994) 'Pragmatic operators'. *The Encyclopaedia of Language and Linguistics*, pp. 3250–3253. Oxford: Pergamon Press.
- Asher, N.** (1993) *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer Academic Publishers.
- Asher, N. and A. Lascarides** (1998) 'The semantics and pragmatics of presupposition'. *Journal of Semantics*, 15(2), pp. 239–300. Available from <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.133.635&rep=rep1&type=pdf>>
- Aysu, S.** (2017) 'The use of discourse markers in the writings of Turkish students of English as a Foreign Language: A corpus-based study'. *Journal of Higher Education and Science*. Available from <http://higheredu-sci.beun.edu.tr/pdf/pdf_HIG_1744.pdf>
- Barker, C. and D. Galasinski** (2001) *Cultural Studies and Discourse Analysis: A Dialogue on Language and Identity*. London: Sage.
- Bazanella, C.** (1990) 'Phatic connectives as intonational cues in contemporary spoken Italian'. *Journal of Pragmatics*, 14(4), pp. 629–647.
- Benz, A., M. Stede and P. Kühnlein** (2012) *Constraints in Discourse 3: Representing and Inferring Discourse Structure*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Blakemore, D.** (1987) *Semantic constraints on relevance*. Oxford: Blackwell.
- Blakemore, D.** (2002) *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.
- Brown, G. and G. Yule** (2012) *Discourse Analysis*. Cambridge: Cambridge University Press.
- Celle, A. and R. Huart** (2007) *Connectives as Discourse Landmarks*. Amsterdam: John Benjamins Publishing Company.
- Council of Europe** (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Chen, C. W. Y.** (2006) 'The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners'. *International Journal of Corpus Linguistics*, 11(1), 113–130.
- Chomsky, N.** (2006) *Language and Mind*. Cambridge: Cambridge University Press.

Crewe, W. (1990) 'The Illogic of Logical Connectives'. *ELT Journal*, 44(4), pp. 316–325.

De Beaugrande, R. and W. U. Dressler (1981) *Introduction to Text Linguistics*. London: Longman.

Dušková, L. et al. (2006) *Mluvnice současné angličtiny na pozadí češtiny*. Praha: Academia.

Fairclough, N. (1995) *Critical Discourse Analysis*. London: Longman.

Fischer, K. (2006) *Approaches to discourse particles*. Amsterdam: Elsevier.

Fraser, B. (1988) 'Types of English discourse markers'. *Acta Linguistica Hungarica*, 38(1–4), pp. 19–33.

Fraser, B. (1990) 'An approach to discourse markers'. *Journal of Pragmatics*, 14, pp. 383–395.

Fraser, B. (1996) 'Pragmatic markers'. *Pragmatics*, 6(2), pp. 167–90. Available from <<https://journals.linguisticsociety.org/elauguage/pragmatics/article/download/418/418-712-1-PB.pdf>>

Fraser, B. (1998) 'Contrastive discourse markers in English', in A. Jucker and Y. Ziv (eds.) *Discourse Markers: Description and Theory*, pp. 301–326. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Fraser, B. (1999) 'What are discourse markers?'. *Journal of Pragmatics*, 31(7), pp. 931–952. Available from <<https://pdfs.semanticscholar.org/df4b/4b4f9a41fced680c30e06dd9db3aff603b2b.pdf>>

Fraser, B. (2009) 'An account of discourse markers'. *International Review of Pragmatics*, 1, pp. 1–28. Available from <<https://www.scribd.com/document/239400620/Fraser-An-Account-of-Discourse-Markers>>

Granger, S., and Tyson, S. (1996) 'Connector usage in the English essay writing of native and non-native EFL speakers of English'. *World Englishes*, 15, 17–27.

Gutwinski, W. (1976) *Cohesion in Literary texts: A Study of Some Grammatical and Lexical Features of English Discourse*. The Hague: Mouton.

Halliday, M. A. K. and R. Hasan (1976) *Cohesion in English*. London: Longman.

Halliday, M. A. K. (1994) *An Introduction to Functional Grammar*. London: Edward Arnold.

Hoey, M. (1991) *Patterns of Lexis in Text*. Oxford: Oxford University Press.

Hunt, K. W. (1996) 'Recent measure in syntactic development'. *Elementary English*, 43, pp. 732–739.

- Irwin, J. W.** (1980) 'The Effects of Explicitness and Clause Order on the Comprehension of Reversible Causal Relations'. *Reading Research Quarterly*, 15(4), pp. 477–488. Available from <http://www.jstor.org/stable/747275?seq=1#page_scan_tab_contents>
- Jaworski, A. and N. Coupland** (2014) *The Discourse Reader*. London: Routledge.
- Johnstone, B.** *Discourse analysis* (2008) London: Blackwell.
- Kwan, Y. H.** (2017) 'The teaching of additive connectives: An evaluation of current English grammar textbooks in Hon Kong'. *The Asian Journal of Applied Linguistics*, 4(2), pp. 161–172.
- Keating, L. C.** (1942) 'The importance of adverbs and connectives in teaching reading'. *The Modern Language Journal*, 26(5), pp. 361–363.
- Lahuerta Martinez, A. C.** (2004) 'Discourse markers in the expository writing of Spanish university students'. *Ibérica*, 8, pp. 63–80. Available from <<http://www.aelfe.org/documents/05-RA-8-Lahuerta.pdf>>
- Leech, G.** (1983) *Principles of Pragmatics*. London: Longman.
- Mann, W. and S. Thompson** (1988) 'Rhetorical Structure Theory: Toward a functional theory of text organization'. *Text*, 8(3), pp. 243–281. Available from <http://www.sfu.ca/rst/05bibliographies/bibs/Mann_Thompson_1988.pdf>
- Marcus, M. et al.** (1993) 'Building a large annotated corpus of English: the Penn Treebank'. *Computational Linguistics*, 19(2), pp. 313–220.
- Miltsakaki, E., R. Prasad, A. Joshi and B. Webber** (2004) *Annotating Discourse Connectives and Their Arguments*. Boston: Association for Computational Linguistics. Available from <<https://www.seas.upenn.edu/~pdtb/papers/frontiers04.pdf>>
- Neuner, J. L.** (1987) 'Cohesive ties and chains in good and poor freshman essays'. *Research in the teaching of English*, 21, pp. 92–103. Available from <<http://www.jstor.org/stable/40171102>>
- Palmer, M., P. Kingsbury and D. Gildea** (2005) 'The Proposition Bank: An Annotated Corpus of Semantic Roles'. *Computational Linguistics*, 31(1), pp. 71–106.
- Pečený, P.** (2017) *The Use of Connectors in the Texts of Non-Native Speakers of Czech*. Thesis supervisor doc. RNDr. Vladimír Petkevič, CSc. Available from <<https://ucjtk.ff.cuni.cz/wpcontent/uploads/sites/57/2017/09/PavelPe%C4%8Den%C3%BD.pdf>>
- Polanyi, L. and M. Van den Berg** (1996) 'Discourse structure and discourse interpretation'. *Proceedings of the Tenth Amsterdam Colloquium, University of Amsterdam*, pp.

- 113-131. Available from
<<https://journals.linguisticsociety.org/proceedings/index.php/BLS/article/view/1263/1047>>
- Prasad, R. et al.** (2007) *The Penn Discourse Treebank 2.0 Annotation Manual*. Philadelphia: Institute for Research in Cognitive Science, University of Pennsylvania. Available from <<https://www.seas.upenn.edu>>
- Redeker, G.** (1991). 'Linguistic markers of discourse structure'. *Linguistics*, 29(6), pp. 1139–1172.
- Rouchota, V.** (1996) 'Discourse connectives: What do they link?'. *UCL Working Papers in Linguistics*, 8.
- Rysová, M.** (2015) *Diskurzní konektory v češtině (Od centra k periferii)*. Thesis supervisor prof. PhDr. Eva Hajičová, DrSc. Available from <<https://is.cuni.cz/webapps/zzp/detail/148691>>
- Rysová, M. and K. Rysová** (2014) 'The Centre and Periphery of Discourse Connectives'. *Pacific Asia Conference on Language, Information and Computation*, 28, pp. 452–459. Available from <<https://aclanthology.info/pdf/Y/Y14/Y14-1052.pdf>>
- Sanders, T. J. M., W. P. M. Spooren and L. G. M. Noordman** (1993) 'Coherence relations in a cognitive theory of discourse representations'. *Cognitive Linguistics*, 4, pp. 93–133.
- Schiffrin, D.** (1987) *Discourse Markers*. Cambridge: Cambridge University Press.
- Schiffrin, D.** (1994) *Approaches to Discourse*. Malden (MA): Blackwell.
- Schiffrin, D., D. Tannen, and H. E. Hamilton** (2001) *The Handbook of Discourse Analysis*. Malden (MA): Blackwell.
- Shaw, P.** (2009) 'Linking adverbials in student and professional writing in literary studies: What makes writing mature'. *Academic Writing. At the Interface of Corpus and Discourse*. Charles, M. et al. (eds.). London: Continuum, pp. 215-235.
- Stubbs, M.** (1983) *Discourse Analysis: The Sociolinguistic Analysis of Natural Language*. Chicago: The University of Chicago Press.
- Tomešová, K.** (2017) *Result/inference discourse connectives in academic texts*. Thesis supervisor doc. PhDr. Markéta Malá, Ph.D. Available from <<https://is.cuni.cz/webapps/zzp/detail/166699>>
- Unger, Ch.** (1996) 'The Scope of Discourse Connectives: Implications for Discourse Organization'. *Journal of Linguistics*, 32(2), pp. 403–438. Available from <<http://www.jstor.org/stable/4176374>>

Urgelles-Coll, M. (2009) *Anyway: A Formal Approach to the Syntax and Semantics of Discourse Markers*. Colchester: University of Essex, Department of Language and Linguistics. Available from <<https://www.essex.ac.uk/linguistics/external/clmt/papers/theses/UrgellesColl09.pdf>>

Urgelles-Coll, M. (2010) *Continuum Studies in Theoretical Linguistics: Syntax and Semantics of Discourse Markers*. London: Continuum International Publishing.

Van Dijk, T. A. (1979) 'Pragmatic connectives'. *Journal of Pragmatics*, 3, 447–456. Amsterdam: North-Holland Publishing Company. Available from <<http://www.discourses.org/OldArticles/Pragmatic%20connectives.pdf>>

Van Dijk, T. A. (2011) *Discourse Studies: A Multidisciplinary Introduction*. Los Angeles: Sage.

Quirk, R. et al. (1985) *A Comprehensive Grammar of the English Language*. London: Longman.

Webber, B. and A. Joshi (1998) 'Anchoring a lexicalized tree adjoining grammar for discourse'. *ACL/COLING Workshop on Discourse Relations and Discourse Markers, Montreal*, pp. 8–92. Available from <<http://aclweb.org/anthology/W98-0315>>

Webber, B., M. Stone, A. Joshi and A. Knott (2003) 'Anaphora and discourse structure'. *Computational Linguistics*, 29, pp. 545–587. Available from <<https://www.eecis.udel.edu/~carberry/CIS-885/Papers/Webber-Anaphora.pdf>>

Witte, S. and L. Faigley (1981) 'Coherence, cohesion and writing quality'. *College composition and communication*, 32, pp. 189–204.

Xue, N. (2005) 'Annotating Discourse Connectives in the Chinese Treebank'. University of Pennsylvania, Department of Computer and Information Science.

Yang, A. W. (2015) 'Cohesive chains and writing quality'. *Word*, 40(1–2), pp. 235–254. Available from <<http://www.tandfonline.com/doi/pdf/10.1080/00437956.1989.11435806>>

Zeyrek, D. and B. Webber (2008) 'A Discourse Resource for Turkish: Annotating Discourse Connectives in the METU Corpus'. In: *The 6th Workshop on Asian Language Resources*.

Zikánová, Šárka, E. Hajičová, B. Vidová-Hladká et al. (2015) *Discourse and Coherence: From the Sentence Structure to Relations in Text*. Praha: Ústav formální a aplikované lingvistiky.

Résumé

Hlavním tématem práce jsou diskurzní konektory jako prostředky textové koherence v psaném projevu nerodilých mluvčích angličtiny. Diskurzní konektory jsou prostředky, které spojují dva úseky textu a vyjadřují mezi nimi sémanticko-pragmatické vztahy. Tyto funkce v angličtině splňují například spojky (*and, but, because*), adverbia (*moreover, however, therefore*) nebo některé víceslovné tvary (*to sum up, the reason is*). Konektory se tak výrazně podílejí na strukturaci textu a jeho celkové koherenci. Cílem práce je podrobný rozbor frekvence a druhů diskurzních konektorů v textech psaných nerodilými mluvčími angličtiny, jejichž jazyková pokročilost odpovídá úrovni B2 ve Společném evropském referenčním rámci pro jazyky. Tímto práce může přispět ke specifikaci úrovně B2 v rámci psaného projevu.

Tématem diskurzních konektorů se v posledních letech zabývá čím dál více lingvistů, a to v různých jazycích, například v arabštině (Al-Saif and Markert, 2010), čínštině (Xue, 2005), turečtině (Zeyrek and Weber, 2008) nebo češtině (Rysová, 2016; Zikánová et al., 2015). V publikacích je tento jazykový jev možné najít pod různými názvy, jako například *discourse markers, discourse particles, pragmatic markers, discourse operators, connectors, pragmatic operators, pragmatic connectives, phatic connectives, semantic conjuncts, conjunctive elements, linking adverbials, conjunctive adverbials*.¹⁸ Přístupů k rozboru diskurzních konektorů existuje mnoho, což poukazuje na složitost zkoumaného jevu. Většina lingvistů se shoduje na tom, že diskurzní konektory spojují nějaké úseky textu. O tom, jaké tyto úseky textu jsou, a jak přesně by se měly diskurzní konektory definovat, však shoda neexistuje.

V teoretické části se práce nejprve zabývá dosavadními přístupy k diskurzu a diskurzním konektorům se zaměřením na výzkum v angličtině. Mezi známé publikace o diskurzních konektorech patří například Halliday a Hasanová (1976), Schiffrinová (1987, 1994 2001), Fraser (1988, 1990, 1996, 1999, 2009) a Blakemoreová (1987, 2002). Ačkoliv se tito autoři v terminologii daného jevu poněkud liší, jejich studie se zaměřují na definici konektorů a funkcí, které v textu plní. Teoretická část práce se tedy soustřeďuje na vyjasnění třech základních pojmů: diskurz, diskurzní vztahy a diskurzní konektory. V rámci definice diskurzu jsou popsány dva základní přístupy – funkční a strukturní (formální), které nabízí dva různé pohledy na podstatu jazyka. Dále se zde vysvětlují pojmy koheze a koherence a jejich vztah k diskurzním konektorům. Kapitola o kohezi a koherenci se z velké části zakládá na díle *Cohesion in English* od Hallidaye a Hasanové (1976), které je přehledem o gramatických a

¹⁸ V české literatuře se většinou používá označení „diskurzní konektor“, proto jsou uvedené termíny v angličtině.

lexikálních prostředcích koheze v textu. Součástí kapitoly o diskurzu jsou i přístupy k vymezení základní jednotky diskurzu. Definice základní jednotky diskurzu používaná v této práci je převzata z korpusu *Penn Discourse Treebank* (PDTB; Prasad et al., 2007). Jedná se o rozsáhlý korpus s anotacemi o diskurzní struktuře a sémantice, který se zaměřuje na analýzu diskurzních vztahů. Tento korpus je i základem pro klasifikaci sémanticko-pragmatických druhů vztahů v textu. Důležitou součástí teoretické části je samotná kapitola o diskurzních konektorech. Předmětem této kapitoly jsou hlavní přístupy k definici konektorů, včetně dvou pohledů v rámci teorie koherence a teorie relevance. Kapitola obsahuje též klasifikaci diskurzních konektorů podle různých autorů a přístupů.

V další sekci se práce zabývá angličtinou nerodilých mluvčích. V rámci této kapitoly je podrobně popsána úroveň B2 podle informací ve Společném rámci v sekcích týkajících se gramatiky, psaného projevu a textové koherence. Jsou zde také shrnuty výsledky novodobých výzkumů o diskurzních konektorech v angličtině nerodilých mluvčích v porovnání s angličtinou rodilých mluvčích. Výsledky těchto studií (Aysu, 2017; Chen, 2006; Granger a Tyson, 1996; Lahuerta Martinez, 2004; Shaw, 2009; Tomešová, 2017) ukazují zvýšený výskyt konektorů v pracích studentů v porovnání s rodilými mluvčími a také častý výskyt chyb na straně nerodilých mluvčích v používání konektorů, převážně na rovině stylistické. Rozbory také naznačují, že jak nerodilí mluvčí, tak rodilí autoři používají menší škálu konektorů opakovaně a dávají přednost počáteční pozici konektoru ve větě. Na základě těchto a předchozích studií, popisu úrovně B2 ve Společném rámci jsme stanovili hypotézu, že studenti budou používat omezený počet konektorů, většina z nich budou jednoslovné primární konektory realizované spojkami, adverbii či předložkovými frázemi, budou se vyskytovat spíše uprostřed věty a budou spojovat verbální argumenty v rámci věty. Předpokládali jsme také, že nejčastěji vyjadřovaným vztahem bude vztah aditivní, a že frekvence konektorů celkově nebude příliš vysoká.

Hlavní součástí práce je pak praktický výzkum diskurzních konektorů v textech psaných nerodilými mluvčími úrovně B2. Materiálem pro výzkum je korpus 161 formálních e-mailů, jejichž obsahem je přihláška na výměnný program Erasmus. Autory textů jsou studenti Vysoké školy ekonomické v Praze, kteří mají angličtinu jako druhý jazyk. Data byla analyzována převážně manuálně – v korpusu textů byly individuálně vyznačeny a anotovány všechny výrazy, které se dají za konektory považovat. Tato metoda byla záměrně zvolena místo automatického vyhledávání, což nám umožnilo vyhledat i komplexnější struktury a sekundární konektory. Jednotlivé body praktické části jsou: frekvence konektorů, jejich realizační formy (morfologické a syntaktické kategorie), druhy konektorů (jednoslovné či víceslovné, primární či sekundární), jejich pozice ve větě, rozsah (mezivětné či vnitrovětné), dále syntaktická povaha

diskurzních argumentů (verbální či nominální) a sémanticko-pragmatické druhy vztahů, které konektory vyjadřují. V rámci jednotlivých podkapitol jsou nasbírané údaje vždy zhodnoceny kvantitativně, vybrané jevy jsou probrány detailně na příkladech ze studentského korpusu textů. Výsledky jsou porovnány s pracovní hypotézou a s dosavadními výzkumy o diskurzních konektorech.

Výsledky práce se s hypotézou v některých ohledech shodují, v některých zase liší. Výběr konektorů používaných autory textů B2 byl sice poměrně široký, ale ukázalo se, že pár základních výrazů, které studenti používali opakovaně, tvořilo většinu výskytů. Nejčastěji se v textech vyskytovala souřadící spojka *and*, dále podřadící spojka *because* a adverbia *so* a *also*. Většina konektorů byly jednoslovné primární konektory realizované spojkami, adverbii a předložkovými frázemi. V textech se vyskytovaly ale i spojovací výrazy realizované celými větami. Tyto spojovací výrazy byly všechny varianty kauzálních frází *The reason is (Důvodem je...)* a *That is why (To je, proč...)*. Většina konektorů se objevovala na začátku argumentu a jejich rozsah byl pouze v rámci věty. Nejčastějším sémanticko-pragmatickým typem vztahu byla expanze, po níž následovaly vztahy kauzální (příčina, důvod). Převaha kauzálních vztahů a vztahů expanze souvisí s frekvencí konektorů *and*, *because*, *so* a *also*. Toto může být však ovlivněno druhem textu, který od mluvčích vyžadoval vysvětlení důvodů, proč se hlásí na výměnný program. Výsledky této práce se shodovaly v dalších aspektech s výsledky novodobých výzkumů o konektorech v textech nerodilých mluvčích. Lze konstatovat, že mluvčí B2 používali konektory nadměrně. Spojka *and* byla často používaná v místech, ve kterých spolu spojované úseky textu nijak nesouvisely. Studenti také často používali více konektorů najednou pro vyjádření stejných vztahů, nebo dokonce vztahů, které si odporovaly. Chyb studentů v užívání konektorů bylo mnoho a týkaly se převážně pozice konektoru ve větě, jejich významu a stylistiky.

Ačkoliv se v textech nerodilých mluvčích objevovalo mnoho nejasností a chyb, a to nejen z hlediska používání diskurzních konektorů, výsledky této práce mohou být využity k bližší specifikaci úrovně B2 ve Společném evropském referenčním rámci, v oblasti psaného jazyka a textové koherence. V současné podobě totiž Společný rámec neobsahuje žádné detailní informace o prostředcích, které nerodilí mluvčí používají pro spojování textu, jsou v něm jen obecné body o tom, zda studenti nějaké prostředky používají či ne. Výsledky této práce tak mohou napomoci při hodnocení jazykové úrovně nerodilých mluvčích pomocí konkrétních jazykových prostředků. Práce také může sloužit jako návrh jazykových oblastí, na které by se nerodilí mluvčí měli soustředit pro vylepšení svého psaného projevu v angličtině.

Appendix

The table below comprises all examples from the texts that were chosen for closer analysis. The number of the text from which individual examples were extracted is in the right column. Due to the extent of the analyzed corpus of texts, the full texts are included only electronically and are accessible in the online depository of the Charles University.

Example No.	Example	Text No.
1	I can speak German, English and Spanish.	47
2	I am studying Business and Management at University of Economics in Czech Republic.	49
3	Dear Mr. Novak, my name is XY and I am writing to you because I heard about your Erasmus offer in Brighton, England.	118
4	I am studying in Prague, where I also live and I am studying first year of University of Economics.	87
5	I would like to thank you for giving such an opportunity for young generation and I hope my dream will come true.	113
6	Thank you for your time and I will be looking forward to hearing from you.	160
7	I would like to travel on my own and lastly I would like to be a part of your institution.	117
8	Firstly, I would like to introduce myself, my name is XY, I am 23 years old and I am studying at the Economics University in Prague, our main study program is based on the Business Administration.	68
9	My name is XY. I am 22 years old. At the moment I am studying Accounting at the University of Economics in Prague.	71
10	I would like to apply for this exchange program because it is interesting for me to get this experience.	93
11	Because I am student of economics and management, I found suitable host university in Paris for me.	94
12	The reason why I want to attend this Erasmus project is mainly because I would like to improve my english or spanish or learn a new language if it is possible.	41

13	I am very interested in this project especially if I will attend your university.	55
14	It could give me worth experience and also help me in my future career.	80
15	I am twenty years old student of Economic University in Prague and exactly I am studying business and law.	97
16	It's been a tough decision for me to choose what would I do in the future during my studies at the grammar school, but after all I found myself in leading people, working on projects, meeting deadline and being very enthusiastic.	100
17	I would love to take the opportunity to attend the Barcelona's University of Economics mainly because of its beautiful city and moreover because of its one of the highest ranking as for the university level of education.	112
18	I am studying at Economic university in Prague, especially I am studying informatic.	152
19	I am studying University of Economics in Prague, exactly social-economic demography.	7
20	At the beginning I would like to write what I am learning now.	27
21	Well, thank you for reading and I hope you will give me a positive response.	105
22	Well, I have already had work experience. (...) Well I think this program can give me experiences, new skills and definitely new friendships.	129
23	Now I am studying at University of Economics in Prague.	5
24	Currently I am studying on the Faculty of International relations on the University of Economics in Prague.	61
25	I have decided to write you in purpose of studying...	27
26	I am writing for purpose, that I would like to apply for exchange program ...	52
27	I am interested in the university because of its history and reputation.	14
28	I am writing to you for I have always wanted to gain much needed experience by studying abroad.	86
29	I have choosed it for its history, profesors that teach at this university and its educational system.	24

30	This text I write you, because I want study in foreign country (for better English knowledge).	7
31	There are some reasons, the first is that I want improve my English skills. I think it's best how I can prove my English skill. The second reason is, that I want be able to speak with anybody in the world. Then I want discover some new cultural and business habits. And the most important reason is that after school I would like work abroad and this is the best way, how I can get some experience.	82
32	My wanted specialization is in Marketing sector, because I'm already working in Prague radio station in this department.	3
33	I'm a Vietnamese but my nanny was Czech.	22
34	I also have finished English course at level B2	14
35	I am interested in your institution because it is the best in what I want, in management.	54
36	The second reason is that Wonderland is the most beautiful country in the world.	54
37	Although I rely not only on academic education, I would be extremely happy to be accepted ...	11
38	Before I started my studying at university, I attended gymnasium.	14
39	CBS has a great reputation as one of the best business schools in EU and the world. So it will be a great contribution to my future success.	18
40	Besides, I have been studying German since my first year in secondary school.	14
41	I am a big dreamer, a creative thinker and a hardworking workaholic. In my past I have been working for companies like Raiffeisen, Vodafone or Nexum. So I think I can apply a lot of my skills and experience in your program.	69
42	I have been studying Spanish at my former high school and also obtained certificate at L1 level. This year I passed my CAE exam as well.	67
43	I have worked in a small corporate in my country, I my work was about planning strategy. I worked on my own projects, too.	73
44	I am contacting you because I am interested in participating at the Erasmus student exchange program.	8

45	I would practice my English a lot and learn new stuff...	14
46	I have always wanted to live in London and gotten to know English culture better ...	95
47	This would help me continue on my studies and to improve my language skills.	2
48	I had to quit them because of health problems.	100
49	I have choosed it for its history, profesors that teach at this university and its educational system.	24
50	I am studying University of Economics in Prague, exactly social-economic demography.	7
51	I study at University of Economics in Prague, exactly business economic and management ...	44
52	Since more experience, meet with new interesting people and other cultures in the world ...	133
53	I would like to study at your university because I am keen on French and I want to improve my French language skills.	70